2012

Structure-function analyses of the Tre1 G protein-coupled receptor

Margaret Marie Pruitt
Iowa State University

Structure-function analyses of the Tre1 G protein-coupled receptor

by

Margaret Marie Pruitt

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Molecular, Cellular and Developmental Biology

Program of Study Committee
Clark R. Coffman, Major Professor
Jeffrey J. Essner
Mary Heather West Greenlee
Monica H. Lamm
Maura McGrail

Iowa State University

Ames, Iowa

2012

TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION

Cell migration is a process that is essential to animals during their entire lifetime. It is critical during animal development when the embryo is undergoing gastrulation, a process that forms a multilayered embryo containing the newly formed germ layers. Additionally, cell migration is important in the immune system, where leukocytes migrate to the area of infection and can either kill the infected cell or activate macrophages. Improper cell migration is also implicated in diseased states such as cancer. Cancer involves metastasizing cells detaching from the primary tumor, migrating through other tissues and eventually colonizing at a second site [1-5]. Cancer is the second most common cause of death in the United States and complications of metastases are the cause of 90% of those deaths [6,7], making it incredibly important to understand the mechanisms and regulation of cell migration. Understanding the mechanisms of cell movements will enable the development of new techniques or approaches to treat cancer and other diseased states, such as multiple sclerosis, caused by improper cell migration.

Primordial germ cell migration in *Drosophila melanogaster* is an excellent model to study cell migration. The movements of Drosophila primordial germ cells are well documented and some of the molecules involved have been identified. The Drosophila primordial germ cells, like in many other animals, are formed at a place distant to the presumptive gonads requiring the primordial germ cells to migrate to their target tissues [8]. In order for the primordial germ cells to properly migrate to the presumptive gonads, the primordial germ cells are required to initiate migration, migrate through the posterior midgut epithelium and the mesoderm, evade or suppress cell death mechanisms, and respond to directional cues. Although some of the molecules involved in Drosophila primordial germ cell migration have been identified, the signaling networks regulating primordial germ cell migration remain incomplete.

Dissertation Organization

In this introductory chapter the following topics will be discussed: germ cell migration in Drosophila, a brief description of germ cell migration in zebrafish and mouse, and an overview of G protein-coupled receptors (GPCRs). The work in this dissertation will focus on the GPCR Trapped in endoderm-1 (Tre1), the roles of

Tre1 in Drosophila germ cell migration, and how Tre1's structure relates to its function. Chapter 2 is a published manuscript in which the mutant allele of *tre1*, *tre1*^{sctt}, is described, as well as the molecular basis for its phenotype. The primary author of this manuscript, Angela Kamps, constructed the transgenic lines and performed half of the germ cell counts. The other half of the germ cell counts, the staining and the statistics were performed for this dissertation. Chapter 3 describes a project whose aim was to dissect the similarity in phenotypes of *tre1*^{sctt} mutant embryos and embryos lacking *wunen* and *wunen2*, two lipid phosphate phosphatases. This project was performed in collaboration with Dr. Martin Schmidt of Des Moines University, who performed the yeast two-hybrid interaction assays. Besides the yeast two-hybrid interaction assays, the work in this chapter was performed for this dissertation. Chapter 4 reports work done to investigate the structural differences between wild-type Tre1 and the mutant protein, Tre1^{sctt}, to better understand the function of Tre1 in primordial germ cell migration. This project was performed in collaboration with Dr. Monica Lamm. Dr. Lamm has served as co-PI on this project and provided the necessary resources and expertise to allow the project to continue. Chapter 5 contains the general conclusions from Chapters 2, 3, and 4 as well as ideas for future directions for each of the projects.

Drosophila Germ Cell Migration

In Drosophila, the primordial germ cells are the first cells to cellularize, forming at the posterior pole of the embryo during stage 3 [9]. After 0-2 rounds of nonsynchronous cell divisions, there are approximately 40 primordial germ cells present at the posterior pole of the embryo [9,10]. The primordial germ cells are left at this extraembryonic location until gastrulation. As the embryo gastrulates, the primordial germ cells adhere to the underlying cells, those of the posterior midgut primordium, and are swept into the developing posterior midgut pocket during germband extension. The primordial germ cells begin their active migration program at stages 9-10. The active migration of the primordial germ cells can be separated into three phases. First, the primordial germ cells initiate migration and move out of the midgut, crossing the midgut epithelium [11-13]. Second, the primordial germ cells migrate dorsally along the midgut surface [14,15]. And lastly, the primordial germ cells migrate into the overlying mesoderm, bilaterally segregate into clusters on either side of the central nervous system (CNS), and move to and align with the somatic gonadal precursor cells. By stage 14, the

primordial germ cells and the somatic gonadal precursor cells have coalesced to form the two embryonic gonads [8,16].

During the first phase of active migration, the primordial germ cells become motile while in the posterior midgut pocket and then proceed to migrate across the midgut epithelium. One of the first genes necessary for active migration at stages 9-10 is the GPCR *tre1*, which is maternally provided and functions in the primordial germ cells. In the complete loss-of-function allele of *tre1*, $tre1^{\Delta EP5}$, the primordial germ cells do not exit the midgut and therefore fail to migrate to the gonads [17]. ΔEP5 was generated by an imprecise excision of EP0496, a P-element located in between *tre1* and the neighboring gene, *Gr5a*. The 5' untranslated regions and the first exons of both *tre1* and *Gr5a* are deleted in ΔEP5 [18,19]. In addition to the role of *tre1* in the transepithelial migration of the primordial germ cells, *tre1* also functions in the polarization and individualization of the primordial germ cells [13]. Tre1 activation causes G protein Gβ, Rho1 and DE-cadherin to change their distribution from the cell periphery to the tail regions of the cells [13]. This rearrangement is thought to aid in the dispersal of the primordial germ cells. In a second loss-of-function allele of *tre1*, $tre1^{sctt}$, the primordial germ cells also fail to migrate properly to the gonads. Similar to $tre1^{\Delta EP5}$, very few germ cells migrate to the gonads [20,21].

The transepithelial migration and dispersal of Drosophila primordial germ cells also requires two lipid phosphate phosphatases, Wunen (Wun) and Wunen2 (Wun2) [22]. Wun and Wun2 are six-pass transmembrane proteins with the phosphatase catalytic domains facing the extracellular environment. *wun* and *wun2* have redundant functions and the two genes are expressed in both the primordial germ cells and somatic tissues [14,15,22]. In *wun* and *wun2* double mutant embryos, lacking *wun* and *wun2* maternally and zygotically, the primordial germ cells fail to exit the posterior midgut during stage 10 [22]. The primordial germ cells in these mutant embryos remain in a tight cluster in the midgut and fail to disperse as individual cells. The lack of migration is not due to an inability to downregulate adhesion between the primordial germ cells as embryos lacking *wun*, *wun2*, and *shotgun* (DE-cadherin) maternally and *wun* and *wun2* zygotically disperse as individual cells but remain within the posterior midgut. The failure to migrate is also not due to a lack in germ cell motility as judged by live imaging [22].

In addition to the functions of *tre1*, *wun* and *wun2* in the first phase of active migration, the posterior midgut structure must change significantly during this time to allow the primordial germ cells to pass through it. The distribution of adherens junctions changes, which leaves noticeable spacing in between the midgut cells, allowing the primordial germ cells to cross the midgut [11,12].

The second phase of active migration is the reorientation of the primordial germ cells on the basal surface of the midgut. The primordial germ cells exit the midgut as individual cells [13] and then migrate to the basal side of the midgut epithelium. This movement is due to the directional cues the primordial germ cells receive from Wun and Wun2 in the somatic cells. At this stage (stage 10 – early stage 11), *wun* and *wun2* RNAs are expressed on the ventral side of the midgut [14,15,23,24]. In wild-type embryos, the primordial germ cells will avoid this area of *wun* and *wun2* expression, and migrate to the basal surface of the midgut. In contrast, in embryos lacking zygotic *wun* and *wun2*, the expression of *wun* and *wun2* on the ventral side of the midgut is absent and there is no reorientation of the primordial germ cells [14,25]. Instead, the primordial germ cells exit the posterior midgut both ventrally and dorsally.

The third and last phase of active migration is the migration of the primordial germ cells into the mesoderm, their separation into two lateral clusters, and their alignment with the somatic gonadal precursor cells. Similar to the migration to the basal surface of the midgut, this phase of primordial germ cell migration is guided by directional cues. After *wun* and *wun2* are expressed on the ventral side of the midgut, the location of *wun* and *wun2* expression changes to the CNS at the midline of the embryo as well as the ectoderm [14,15,23,24]. Interestingly, the areas of *wun* and *wun2* expression flank the migratory route of the primordial germ cells. From their position on the basal surface of the midgut, the primordial germ cells next migrate into the overlying mesoderm. Once in the mesoderm, the primordial germ cells split into two clusters on either side of the *wun* and *wun2* expressing midline and continue to migrate to the developing somatic gonads. When an embryo is missing *wun* and *wun2* zygotically, primordial germ cell migration is disrupted. The primordial germ cells exit the posterior midgut, however, the primordial germ cells fail to laterally sort due to a lack of *wun* and *wun2* expression in the midline. Few primordial germ cells migrate properly to the gonads. In order to see this migration phenotype, both *wun* and *wun2* must be absent, which suggests redundancy of function [14,15,23,26-28].

Overexpression experiments were performed to determine if *wun* and *wun2* act as repellants and if *wun* and *wun2* catalytic activity is necessary for this function. When either *wun* or *wun2* is overexpressed in the mesoderm, usually an attractive or passive tissue where *wun* and *wun2* are not normally expressed, the primordial germ cells fail to enter the mesoderm and instead stay on the midgut. The catalytic activity of Wun2 is necessary for this phenotype [14,15]. Additionally, overexpressing *wun2* in the CNS in a *wun wun2* mutant background partially restores lateral segregation of the primordial germ cells [28]. The current model for Wun and Wun2 activity is that Wun and Wun2 hydrolyze, and thus destroy, a phospholipid or phospholipid-modified attractant. According to this model, in tissues expressing Wun and Wun2, there is a very low concentration of this phospholipid attractant. The low attractant concentration causes the primordial germ cells to avoid these areas, instead migrating to areas of higher attractant concentration [14,27,28].

As the primordial germ cells are responding to the signals from Wun and Wun2 during the final phase of active migration, the primordial germ cells must also respond to attractive signals. In Drosophila, the primordial germ cells are attracted to the gonadal mesoderm by a signal dependent on 3-hydroxy-3-methylglutaryl coenzymeA reductase (HMGCR) function. HMGCR catalyzes the synthesis of mevalonate, a precursor to isoprenoids and cholesterol. *hmgcr* is expressed in the somatic gonadal mesoderm, and ectopic expression of *hmgcr* is sufficient to attract primordial germ cells to ectopic locations [29]. The somatic gonadal mesoderm is located in parasegments 10, 11 and 12, an area dorsal to where the primordial germ cells are situated at stages 10-11 [30,31]. As further evidence that *hmgcr* provides an attractive signal, two enzymes that act downstream of *hmgcr* in the isoprenoid branch of the pathway (*fpps* which encodes a Farnesyl diphosphate synthase and *quemao* which encodes a Geranylgeranyl diphosphate synthase) have germ cell migration phenotypes similar to *hmgcr* loss-of-function and are sufficient to attract primordial germ cells to ectopic locations when overexpressed. Additionally, embryos mutant for the enzyme *βGGT1*, which encodes Geranylgeranyl transferase type 1, also display germ cell migration defects. The involvement of these enzymes is consistent with the model that the primordial germ cell attractant is geranylgeranylated [32]. If the attractant was geranylgeranylated, it would need to be exported from the somatic gonadal precursor cells in order for it to attract the primordial germ cells. Ricardo and Lehmann (2009) showed that an ABC transporter encoded by

multidrug resistance 49 (*mdr49*) acts in the mesoderm and is required to attract the primordial germ cells. Epistasis experiments were performed to show MDR49 transports a geranylgeranylated attractant [33].

There is also evidence suggesting a role of Hedgehog (Hh) signaling in germ cell attraction. *hh* is expressed in the somatic gonadal precursor cells and ectopic expression of *hh* is sufficient to cause primordial germ cell migration to that ectopic location. Additionally, mutants of proteins involved in *hh* signaling (*smoothened*, *fused*, *patched*, and *protein kinase A*) show primordial germ cell migration defects [34]. Evidence also connects the *hmgcr* pathway to the *hh* signaling pathway. It has been suggested that *hmgcr* functions in *hh* signaling by geranylating Gγ1, which is linked to germ cell attraction as geranylated Gγ1 enhances the release of the Hh signal [35,36]. Although results indicate a role of Hh signaling in germ cell attraction, whether or not Hh signaling is involved is unclear. Findings by the Lehmann Lab argue there is no role of Hh signaling in germ cell attraction [33,37].

Throughout the active migration of the Drosophila primordial germ cells, the cells not only respond to directional cues, but also respond to death and survival signals. Initially, approximately 40 primordial germ cells form at the posterior pole, yet only 50% of the primordial germ cells, on average, are incorporated into the gonads. Lineage analysis and transplantation experiments have shown that the primordial germ cells that are not incorporated into the gonad die rather than transdifferentiating [9,38,39]. Surprisingly, not much is known about the mechanism of primordial germ cell death. Work performed by Yamada et al. demonstrated that 50% of the primordial germ cells die during stages 10-12, prior to coalescence of the gonads. Two genes, *outsiders* and *p53*, were found to be involved in this process [40]. The *outsiders* alleles were generated in an EMS mutagenesis screen [20], and the gene encodes a putative monocarboxylate transporter. Primordial germ cells migrate to the gonads properly in *outsiders* mutants, but many primordial germ cells are left ectopic to the gonads and fail to die. Germ cell death is also disrupted in loss-of-function *p53* alleles, which have a similar phenotype to *outsiders*. It was shown using genetics that *outsiders* and *p53* may function in the same or closely linked pathways [40]. Even though the timing of when the primordial germ cells die has been established, how the primordial germ cells die is unknown.

The other two genes known to be involved in germ cell death are the lipid phosphate phosphatases, *wun* and *wun2*. *wun* and *wun2* involvement in germ cell death occurs in two ways. First, a competition exists

between the soma and the germline for a lipid phosphate substrate. The current model is that one function of *wun* and *wun2* in the primordial germ cells is to take up a lipid phosphate that is required for primordial germ cell survival, and this same lipid phosphate is taken up by *wun* and *wun2* in the soma to form a phospholipid gradient used by the migrating primordial germ cells. If the levels of *wun* and *wun2* change in the soma so that there is more *wun* and *wun2* expressed in somatic tissues, the primordial germ cells are unable to take up enough of this "survival" signal and die [26,27]. Experiments using either an antibody against cleaved caspase 3, overexpression of the caspase inhibitor p35 or TUNEL staining suggest that this primordial germ cell death occurs in a caspase-independent manner [26,27]. Second, *wun* and *wun2* are involved in germ cell death at the midline. After the primordial germ cells exit the midgut, they sort into two lateral clusters and continue migration to the somatic gonad. Any primordial germ cells left in the midline, those that do not laterally sort, die in a way that is dependent on *wun* and *wun2*. In mutants lacking both *wun* and *wun2* zygotically, primordial germ cells do not laterally sort and the primordial germ cells left in the middle of the embryo do not die. This death that is dependent on *wun* and *wun2* does not appear to be by apoptosis [28].

Vertebrate Germ Cell Migration

While primordial germ cell migration has been studied extensively in Drosophila, substantial work has also been performed in zebrafish and mouse model systems. Similar to Drosophila, the primordial germ cells in zebrafish are specified in a location far from the developing gonad. The main molecules involved in zebrafish primordial germ cell migration are the GPCR-ligand pair chemokine (CXC motif) receptor 4b (CXCR4b) and Stromal-derived factor 1 (SDF1a) [41,42]. CXCR4b is expressed in the primordial germ cells and the dynamic expression of SDF1a in somatic tissue leads to proper migration first to an intermediate target and next to the final target, the gonads [43].

Again, like in Drosophila and zebrafish, the mouse primordial germ cells need to travel a long distance from the place where they first differentiate, the posterior primitive streak, to the genital ridges. Similar to Drosophila primordial germ cell migration, the primordial germ cells in mouse migrate through hindgut tissue to the mesoderm. In the mesoderm, the primordial germ cells separate into two lateral clusters and migrate to the genital ridges [44]. The motility and survival of mouse primordial germ cells is regulated by the receptor

tyrosine kinase c-kit and its ligand Steel, and the primordial germ cells express c-kit and the somatic cells express Steel throughout all stages of migration [45]. Directional cues during mouse primordial germ cell migration to the genital ridges are the function of CXCR4 and SDF1. CXCR4 is expressed in the migrating primordial germ cells and SDF1 is expressed in the target tissue, the genital ridges [46,47].

G Protein-Coupled Receptors

GPCRs are the largest family of cell surface receptors, with the human GPCR superfamily consisting of more than 800 members [48-50]. GPCRs are critical for many biological functions such as cell migration, proliferation and differentiation, and GPCR dysfunction can lead to diseased states such as cancer or atherosclerosis. The prevalence of GPCR involvement in many diseased states has made GPCRs of particular interest to the drug industry [48], with approximately one-third of approved drugs targeting GPCRs [50,51].

Although there are many different genes encoding for GPCRs, and GPCRs function in a variety of processes, all GPCRs share a common structure. GPCRs thread through membranes seven times and contain three extracellular loops and three intracellular loops. There are five main families of human GPCRs (Rhodopsin, Secretin, Adhesion, Glutamate, Frizzled/Taste2) [48-50], and this classification holds true for GPCRs in other bilateral species [52]. The Rhodopsin family is the largest family of GPCRs, and includes the Tre1 GPCR involved in Drosophila primordial germ cell migration [13,17].

Within Rhodopsin family GPCRs there are four groups (α, β, γ and δ) and many sub-branches [48]. There are seven conserved sequence motifs in Rhodopsin family GPCRs falling within the transmembrane (TM) regions. Three of these motifs (DRY in TM3, CWxP in TM6 and NPxxY in TM7) are thought to act as micro-switches in the activation mechanism of GPCRs [50,53]. A great deal of research has been done on the DRY motif and it is thought to be involved in holding the receptor in either an active or inactive conformation depending on the receptor and environmental context [54-65]. While GPCRs are a large class of proteins that share a common structure, beyond the seven TM regions GPCRs have very diverse structures. For example, the human β_2-adrenergic receptor is part of the α branch of Rhodopsin family GPCRs, and it shares only 25% sequence identity to human CXCR4 that is part of the γ branch of Rhodopsin family GPCRs [66].

In general, GPCRs function to receive extracellular cues and transform the cues into intracellular

responses such as cell migration, proliferation, survival and differentiation. Ligands for GPCRs are both diverse in size and chemical properties. Ligands for GPCRs can be amines, amino acids, photons, ions, lipids, peptides, proteins, neurotransmitters or hormones [50,67]. Canonical GPCR signaling involves a receptor, a G protein and an effector. Specifically, a ligand binds to the extracellular or TM region of a GPCR, which results in the receptor changing conformation. The conformational change of the GPCR activates the cytoplasmically attached heterotrimeric G protein. In the unstimulated state, the G protein is a complex consisting of three subunits, α, β and γ, with a GDP bound to the α subunit. Upon activation of the GPCR, the GPCR acts as a guanine nucleotide exchange factor, and the α subunit releases GDP, allowing GTP to bind. The GDP to GTP transition causes the G protein trimer to dissociate into two units, α and βγ. Both α and βγ are now activated and act to mediate cellular responses. G proteins can signal through various means such as regulating production of cyclic AMP, activating phospholipase C-β, or acting to regulate ion channels [68]. GPCRs can contain multiple active sites for different ligands and exist in multiple conformations, not simply as a bimodal on and off switch [69]. In addition, some GPCRs do not signal through G proteins, instead signaling through molecules such as G protein-coupled receptor kinases and β-arrestins [70,71].

One important response mediated by GPCRs is cell migration, which is an important process in development and disease progression. A specific way GPCR signaling leads to cell migration is by initiating signaling cascades that regulate cell shape changes and the formation of a leading edge. GPCR activation can cause actin remodeling and thus aid in cell migration [72]. Activation of the G protein α_i can promote the activation of Rho, which regulates aspects of actin remodeling. Rho proteins are small GTP-binding proteins that can mediate the reorganization of the cytoskeleton, [73], and Rho GTPases can regulate WASP/WAVE proteins that are important for activating the Arp2/3 complex to promote the formation of new actin filaments [74].

Many diseases are caused by loss-of-function mutations in different types of GPCRs. Loss-of-function mutations in cone opsins can lead to color blindness, loss-of-function mutations in Endothelin-B cause Hirschsprung disease, and mutations to gonadotropin-releasing hormone receptor can cause central hypogonadism [68]. Diseased states are also caused by errant expression of GPCRs. A receptor for chemokines that is involved in directional migration, CXCR4, is expressed in high amounts in breast cancer

cells and this increased expression directs invasive responses [75]. A similar result is seen with the bioactive lipid lysophosphatidic acid (LPA) that signals through six GPCRs, LPA1-LPA6. Elevated LPA signaling or errant expression of a receptor can lead to cancer cell proliferation, migration and invasion as well as angiogenesis [76,77].

The involvement of GPCRs in cancer can be reminiscent of the roles GPCRs have in normal development. The chemokine receptor CXCR4 and its ligand SDF-1 are involved in chemotaxis and cell proliferation as well as tumor progression, angiogenesis, metastasis and survival [78]. This receptor-ligand pair is involved in many different types of cancer such as breast, pancreatic and prostate cancers [77]. Although the type of cancer varies, the tumor cells upregulate CXCR4 and the secondary site expresses SDF-1 [67]. GPCRs are also involved in cancer by the transactivation of epidermal growth factor receptor, EGFR. Activation of a GPCR by its ligand can cause a transactivation of EGFR, which in turn leads to intracellular signaling cascades that are important for normal development as well as promoting cancer progression. For example, in breast cancer, the GPCR protease-activated receptor 1 (PAR1) continually activates EGFR and ErbB signaling which promotes breast carcinoma cell invasion [77,79].

References

1. Franz CM, Jones GE, Ridley AJ (2002) Cell migration in development and disease. Dev Cell 2: 153-158.
2. Friedl P, Gilmour D (2009) Collective cell migration in morphogenesis, regeneration and cancer. Nat Rev Mol Cell Biol 10: 445-457.
3. Horwitz R, Webb D (2003) Cell migration. Curr Biol 13: R756-759.
4. Lauffenburger D, Horwitz A (1996) Cell migration - a physically integrated molecular process. Cell 84: 1-11.
5. Ridley AJ, Schwartz M, Burridge K, Firtel R, Ginsberg M, et al. (2003) Cell Migration: Integrating Signals from Front to Back. SCIENCE 302: 1704-1709.
6. ACS (2009) Cancer Facts & Figures 2009. American Cancer Society.
7. Gupta GP, Massagué J (2006) Cancer metastasis: building a framework. Cell 127: 679-695.
8. Kunwar P, Siekhaus D, Lehmann R (2006) In Vivo Migration: A Germ Cell Perspective. Annual Reviews of Cell and Developmental Biology 22: 237-265.
9. Sonnenblick BP (1941) Germ cell movements and sex differentiation of the gonads in the Drosophila embryo. PNAS 27: 484-489.
10. Williamson A, Lehmann R (1996) Germ cell development in Drosophila. Annu Rev Cell Dev Biol 12: 365-391.
11. Callaini G, Giovanna M, Riparbelli, Dallai R (1995) Pole cell migration through the gut wall of the Drosophila embryo - analysis of cell interactions. Developmental Biology 170: 365-375.
12. Jaglarz M, Howard K (1995) The active migration of Drosophila PGC. Development 121: 3495-3503.
13. Kunwar PS, Sano H, Renault AD, Barbosa V, Fuse N, et al. (2008) Tre1 GPCR initiates germ cell transepithelial migration by regulating Drosophila melanogaster E-cadherin. J Cell Biol 183: 1-12.
14. Starz-Gaiano M, Cho N, Forbes A, Lehmann R (2001) Spatially restricted activity of a Drosophila lipid phosphatase guides migrating germ cells. Development 128: 983-991.

15. Zhang N, Zhang J, Purcell K, Cheng Y, Howard K (1997) The Drosophila protein Wunen repels migrating germ cells. Nature 385: 64-67.
16. Campos-Ortega J, Hartenstein V (1997) The Embryonic Development of *Drosophila melanogaste*r. New York: Springer-Verlag. 405 p.
17. Kunwar P, Starz-Gaiano M, Bainton R, Heberlein U, Lehmann R (2003) Tre1, a GPCR, directs transepithelial migration of Drosophila germ cells. Plos Biol 1: 372-384.
18. Dahanukar A, Foster K, van der Goes van Naters WM, Carlson JR (2001) A Gr receptor is required for response to the sugar trehalose in taste neurons of Drosophila. Nat Neurosci 4: 1182-1186.
19. Ueno K, Ohta M, Morita H, Mikuni Y, Nakajima S, et al. (2001) Trehalose sensitivity in Drosophila correlates with mutations in and expression of the gustatory receptor gene Gr5a. Current biology 11: 1451-1455.
20. Coffman CR, Strohm R, Oakley F, Yamada Y, Przychodzin D, et al. (2002) Identification of X-linked genes required for migration and programmed cell death of Drosophila melanogaster germ cells. Genetics 162: 273-284.
21. Kamps AR, Pruitt MM, Herriges JC, Coffman CR (2010) An Evolutionarily Conserved Arginine is Essential for Tre1 G Protein-Coupled Receptor Function During Germ Cell Migration in Drosophila melanogaster. PLoS ONE 5: e11839.
22. Renault AD, Kunwar PS, Lehmann R (2010) Lipid phosphate phosphatase activity regulates dispersal and bilateral sorting of embryonic germ cells in Drosophila. Development 137: 1815-1823.
23. Zhang N, Zhang J, Cheng Y, Howard K (1996) Identification and genetic analysis of wunen, a gene guiding Drosophila melanogaster germ cell migration. Genetics 143: 1231-1241.
24. Renault AD, Starz-Gaiano M, Lehmann R (2002) Metabolism of sphingosine 1-phosphate and lysophosphatidic acid: a genome wide analysis of gene expression in Drosophila. Mechanisms of Development 119 Suppl 1: S293-301.
25. Zhang N, Zhang J, Purcell K, Cheng Y, Howard K (1997) The Drosophila protein Wunen repels migrating germ cells. Nature 385: 4.
26. Hanyu-Nakamura K, Kobayashi S, Nakamura A (2004) Germ cell-autonomous Wunen2 is required for germline development in Drosophila embryos. Development 131: 4545-4553.
27. Renault AD, Sigal Y, Morris A, Lehmann R (2004) Soma-Germ Line Competition for Lipid Phosphate Uptake Regulates Germ Cell Migration and Survival. SCIENCE 305: 1963-1966.
28. Sano H, Renault A, Lehmann R (2005) Control of lateral migration and germ cell elimination by the Drosophila melanogaster lipid phosphate phosphatases Wunen and Wunen 2. The Journal of Cell Biology 171: 675-683.
29. Van Doren M, Broihier HT, Moore LA, Lehmann R (1998) HMG-CoA reductase guides migrating primordial germ cells. Nature 396: 466-469.
30. Boyle M, Bonini N, DiNardo S (1997) Expression and function of clift in the development of somatic gonadal precursors within the Drosophila mesoderm. Development 124: 971-982.
31. Broihier HT, Moore LA, Van Doren M, Newman S, Lehmann R (1998) zfh-1 is required for germ cell migration and gonadal mesoderm development in Drosophila. Development 125: 655-666.
32. Santos AC, Lehmann R (2004) Isoprenoids control germ cell migration downstream of HMGCoA reductase. Developmental Cell 6: 283-293.
33. Ricardo S, Lehmann R (2009) An ABC Transporter Controls Export of a Drosophila Germ Cell Attractant. SCIENCE 323: 943-946.
34. Deshpande G, Swanhart L, Chiang P, Schedl P (2001) Hedgehog signaling in germ cell migration. Cell 106: 756-769.
35. Deshpande G, Godishala A, Schedl P (2009) Gγ1, a Downstream Target for the hmgcr-Isoprenoid Biosynthetic Pathway, Is Required for Releasing the Hedgehog Ligand and Directing Germ Cell Migration. PLoS Genet 5: e1000333.
36. Deshpande G, Schedl P (2005) HMGCoA reductase Potentiates hedgehog Signaling in Drosophila melanogaster. Developmental Cell 9: 629-638.
37. Renault AD, Ricardo S, Kunwar PS, Santos A, Starz-Gaiano M, et al. (2009) Hedgehog does not guide migrating Drosophila germ cells. Dev Biol 328: 355-362.

38. Technau GM, and Campos-Ortega, J.A. (1986) Lineage analysis of transplanted individual cells in embryos of Drosophila melanogaster. III. Commitment and proliferative capabilities of pole cells and midgut progenitors. Roux's Arch Developmental Biology 195: 489-498.
39. Underwood E, Caulton J, Allis C, Mahowald A (1980) Developmental fate of pole cells in Drosophila melanogaster. Dev Biol 77: 303-314.
40. Yamada Y, Davis KD, Coffman CR (2008) Programmed cell death of primordial germ cells in Drosophila is regulated by p53 and the Outsiders monocarboxylate transporter. Development 135: 207-216.
41. Doitsidou M, Reichman-Fried M, Stebler J, Koprunner M, Dorries J, et al. (2002) Guidance of primordial germ cell migration by the chemokine SDF-1. Cell 111: 647-659.
42. Knaut H, Werz C, Geisler R, Consortium TTS, Nüsslein-Volhard C (2003) A zebrafish homologue of the chemokine receptor Cxcr4 is a germ-cell guidance receptor. Nature 421: 279-282.
43. Reichman-Fried M, Minina S, Raz E (2004) Autonomous Modes of Behavior in Primordial Germ Cell Migration. Developmental Cell 6: 589-596.
44. Molyneaux KA, Stallock J, Schaible K, Wylie C (2001) Time-lapse analysis of living mouse germ cell migration. Developmental Biology 240: 488-498.
45. Gu Y, Runyan C, Shoemaker A, Surani A, Wylie C (2009) Steel factor controls primordial germ cell survival and motility from the time of their specification in the allantois, and provides a continuous niche throughout their migration. Development 136: 1295-1303.
46. Molyneaux KA, Zinszer H, Kunwar P, Schaible K, Stebler J, et al. (2003) The chemokine SDF1/CXCL12 and its receptor CXCR4 regulate mouse germ cell migration and survival. Development 130: 4279-4286.
47. Ara T, Nakamura Y, Egawa T, Sugiyama T, Abe K, et al. (2003) Impaired colonization of the gonads by primordial germ cells in mice lacking a chemokine, stromal cell-derived factor-1 (SDF-1). Proc Natl Acad Sci USA 100: 5319-5323.
48. Fredriksson R, Lagerström MC, Lundin L-G, Schiöth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Molecular pharmacology 63: 1256-1272.
49. Gloriam DE, Fredriksson R, Schiöth HB (2007) The G protein-coupled receptor subset of the rat genome. BMC Genomics 8: 338.
50. Lagerström MC, Schiöth HB (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. Nature Reviews Drug Discovery 7: 339-357.
51. Kimple AJ, Bosch DE, Giguere PM, Siderovski DP (2011) Regulators of G-Protein Signaling and Their G Substrates: Promises and Challenges in Their Use as Drug Discovery Targets. Pharmacological Reviews 63: 728-749.
52. Fredriksson R, Schiöth HB (2005) The repertoire of G-protein-coupled receptors in fully sequenced genomes. Molecular pharmacology 67: 1414-1425.
53. Nygaard R, Frimurer TM, Holst B, Rosenkilde MM, Schwartz TW (2009) Ligand binding and micro-switches in 7TM receptor structures. Trends in Pharmacological Sciences 30: 249-259.
54. Aizaki Y, Maruyama K, Nakano-Tetsuka M, Saito Y (2009) Distinct roles of the DRY motif in rat melanin-concentrating hormone receptor 1 in signaling control. Peptides 30: 974-981.
55. Ballesteros J, Kitanovic S, Guarnieri F, Davies P, Fromme BJ, et al. (1998) Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. J Biol Chem 273: 10445-10453.
56. Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, et al. (2001) Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. The Journal of Biological Chemistry 276: 29171-29177.
57. Flanagan CA (2005) A GPCR that is not "DRY". Mol Pharmacol 68: 1-3.
58. Shapiro DA, Kristiansen K, Weiner DM, Kroeze WK, Roth BL (2002) Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. The Journal of Biological Chemistry 277: 11441-11449.
59. Vogel R, Mahalingam M, Ludeke S, Huber T, Siebert F, et al. (2008) Functional Role of the "Ionic Lock" - An Interhelical Hydrogen-Bond Network in Family A Heptahelical Receptors. Journal of Molecular Biology 360: 648-655.

60. Zhu SZ, Wang SZ, Hu J, el-Fakahany EE (1994) An arginine residue conserved in most G protein-coupled receptors is essential for the function of the m1 muscarinic receptor. Molecular pharmacology 45: 517-523.
61. Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. EMBO J 15: 3566-3578.
62. Capra V, Veltri A, Foglia C, Crimaldi L, Habib A, et al. (2004) Mutational analysis of the highly conserved ERY motif of the thromboxane A2 receptor: alternative role in G protein-coupled receptor signaling. Molecular pharmacology 66: 880-889.
63. Greasley PJ, Fanelli F, Rossier O, Abuin L, Cotecchia S (2002) Mutagenesis and modelling of the alpha(1b)-adrenergic receptor highlight the role of the helix 3/helix 6 interface in receptor activation. Molecular pharmacology 61: 1025-1032.
64. Scheer A, Costa T, Fanelli F, DeBenedetti P, Mhaouty-Kodja S, et al. (2000) Mutational Analysis of the Highly Conserved Arginine within the Glu/Asp-Arg-Tyr Motif of the alpha1b-Adrenergic Receptor: Effects on Receptor Isomerization and Activation. Molecular pharmacology 57: 219-231.
65. Greasley PJ, Fanelli F, Scheer A, Abuin L, Nenniger-Tosato M, et al. (2001) Mutational and Computational Analysis of the alpha1b-Adrenergic Receptor: Involvement of Basic and Hydrophobic Residues in Receptor Activation and G Protein Coupling. The Journal of Biological Chemistry 276: 46485-46494.
66. Katritch V, Cherezov V, Stevens RC (2011) Diversity and modularity of G protein-coupled receptor structures. Trends in Pharmacological Sciences: 1-11.
67. Dorsam RT, Gutkind JS (2007) G-protein-coupled receptors and cancer. Nat Rev Cancer 7: 79-94.
68. Spiegel AM, Weinstein LS (2004) Inherited diseases involving G proteins and G protein-coupled receptors. Annual review of medicine 55: 27-39.
69. Huber T, Sakmar TP (2011) Escaping the flatlands: new approaches for studying the dynamic assembly and activation of GPCR signaling complexes. Trends in Pharmacological Sciences 32: 410-419.
70. Hall RA, Premont RT, Lefkowitz RJ (1999) Heptahelical receptor signaling: beyond the G protein paradigm. The Journal of Cell Biology 145: 927-932.
71. Rajagopal K, Lefkowitz RJ, Rockman HA (2005) When 7 transmembrane receptors are not G protein-coupled receptors. Journal of Clinical Investigation 115: 2971-2974.
72. Cotton M, Claing A (2009) G protein-coupled receptors stimulation and the control of cell migration. Cellular Signalling 21: 1045-1053.
73. Etienne-Manneville S, Hall A (2002) Rho GTPases in cell biology. Nature 420: 629-635.
74. Millard TH, Sharp SJ, Machesky LM (2004) Signaling to actin assembly via the WASP (Wiskott-Aldrich syndrom protein)-family proteins and the Arp2/3 complex. The Biochemical journal 380: 1-17.
75. Muller A, Homey B, Soto H, Ge N, Catron D, et al. (2001) Involvement of chemokine receptors in breast cancer metastasis. Nature 410: 50-56.
76. Chen M, Towers LN, O'Connor KL (2006) LPA2 (EDG4) mediates Rho-dependent chemotaxis with lower efficacy than LPA1 (EDG2) in breast carcinoma cells. American Journal of Physiology - Cell Physiology 292: C1927-C1933.
77. Lappano R, Maggiolini M (2011) G protein-coupled receptors: novel targets for drug discovery in cancer. Nature Reviews Drug Discovery 10: 47-60.
78. Teicher BA, Fricker SP (2010) CXCL12 (SDF-1)/CXCR4 Pathway in Cancer. Clinical Cancer Research 16: 2927-2931.
79. Arora P, Cuevas BD, Russo A, Johnson GL, Trejo J (2008) Persistent transactivation of EGFR and ErbB2/HER2 by protease-activated receptor-1 promotes breast carcinoma cell invasion. Oncogene 27: 4434-4445.

CHAPTER 2: AN EVOLUTIONARILY CONSERVED ARGININE IS ESSENTIAL FOR TRE1 G PROTEIN-COUPLED RECEPTOR FUNCTION DURING GERM CELL MIGRATION IN *DROSOPHILA MELANOGASTER*

A paper published in *PLoS ONE*

Angela R. Kamps, Margaret M. Pruitt, John C. Herriges, and Clark R. Coffman

Abstract

G protein-coupled receptors (GPCRs) play central roles in mediating cellular responses to environmental signals leading to changes in cell physiology and behaviors, including cell migration. Numerous clinical pathologies including metastasis, an invasive form of cell migration, have been linked to abnormal GPCR signaling. While the structures of some GPCRs have been defined, the *in vivo* roles of conserved amino acid residues and their relationships to receptor function are not fully understood. Trapped in endoderm 1 (Tre1) is an orphan receptor of the rhodopsin class that is necessary for primordial germ cell migration in *Drosophila melanogaster* embryos. In this study, we employ molecular genetic approaches to identify residues in Tre1 that are critical to its functions in germ cell migration. First, we show that the previously reported *scattershot* mutation is an allele of *tre1*. The *scattershot* allele results in an in-frame deletion of 8 amino acids at the junction of the third transmembrane domain and the second intracellular loop of Tre1 that dramatically impairs the function of this GPCR in germ cell migration. To further refine the molecular basis for this phenotype, we assayed the effects of single amino acid substitutions in transgenic animals and determined that the arginine within the evolutionarily conserved E/N/DRY motif is critical for receptor function in mediating germ cell migration within an intact developing embryo. These structure-function studies of GPCR signaling in native contexts will inform future studies into the basic biology of this large and clinically important family of receptors.

Introduction

Signaling mediated by G protein-coupled receptors (GPCRs) facilitates the transmission of extracellular environmental cues into a cell, regulating a myriad of cellular responses and signaling cascades

involved in development, homeostasis, and disease states. GPCRs represent the largest class of cell surface receptors and include over 800 different receptors in humans [1, 2]. Abnormal GPCR function contributes to the onset of many pathologies including cancer, vascular, and neurodegenerative diseases. For this reason, GPCRs are one of the most common targets for pharmaceutical intervention in the treatment of human disease [3-5].

The rhodopsin family shares the characteristic seven transmembrane structure of GPCRs. Additionally, a highly conserved E/N/DRY (Glutamic Acid/Asparagine/Aspartic Acid-Arginine-Tyrosine) motif is found at the junction of the third transmembrane domain and second intracellular loop of these signal transduction molecules. The arginine of this triplet is a hallmark of these receptors and is conserved in 96% of rhodopsin family GPCRs [6, 7]. Substitution of this residue in tissue culture cells results in disruption of receptor signaling [8-10].

The well-studied chemokine receptors, including CXCR4, are among the GPCRs with the conserved E/N/DRY motif. CXCR4 receptors have established roles in tumor metastasis [11]. They have also been shown to mediate germ cell migration in zebrafish, chickens, and mammals [12-17]. In *Drosophila melanogaster* germ cells, the GPCR Trapped in endoderm 1 (Tre1) appears to be a functional analog to CXCR4 in vertebrates, playing a critical role in the migration of primordial germ cells [18-21].

The *scattershot* (*sctt*) mutation results in severe disruption of germ cell migration [22]. Here we demonstrate that *sctt* causes mis-splicing of the *tre1* transcript, resulting in an in-frame deletion of 8 codons, including two encoding the conserved arginine and tyrosine of the E/N/DRY motif. The *sctt* mutant provided a unique opportunity to perform a structure-function analysis at single amino acid resolution of this evolutionarily conserved motif within an intact developing organism. We more fully characterized the *sctt* mutant phenotype and performed a transgenic rescue analysis that demonstrates that the defects in germ cell migration observed in *sctt* mutants can be rescued by Tre1 constructs containing the arginine of the conserved E/N/DRY motif. Substitution of the arginine with an alanine without altering the other 7 amino acids results in a loss-of-function phenotype that does not rescue *sctt* mutants. This provides evidence that the arginine plays a critical role in maintaining the signaling function of this GPCR.

Materials and Methods

Fly Stocks

The *sctt* allele was generated in an EMS mutagenesis screen [22]. The T^+G^+ transgenic line was kindly provided by John Carlson [27, 46]. w^{1118}, P{w^+, *fat facets-lacZ*}, the unmutagenized parental strain of *sctt*, was used as a wild-type control [47]. Balancer containing stocks were obtained from the Bloomington Drosophila Stock Center.

Embryo Collections

Embryos were collected on standard apple juice agar plates and aged to stages 15-16 at 25 ºC [48]. Embryos were harvested and dechorionated using 50% bleach.

Whole Mount Antibody Staining

Immunostaining was performed according to standard procedures [49]. The primary antibody used was chicken anti-Vasa, (1:10,000, a gift from Ken Howard). The secondary antibody was biotinylated anti-chicken IgG (1:500) (Vector Laboratories). Antibody detection was performed using Vector Laboratories ABC Elite Kits with 3,3'-diaminobenzidine tetrahydrochloride as a substrate.

β-Galactosidase Staining

β-Galactosidase staining of embryos followed published procedures [50, 51]. Staining was performed with a 0.08% X-Gal solution for 2 hours at 37 ºC.

Germ Cell Counts

Germ cells were labeled using either a β-Galactosidase assay or anti-Vasa antibody and counted using a differential interference contrast microscope. Both methods detect similar numbers of primordial germ cells [52]. Embryos were aged to stages 15-16. Staging was confirmed using embryo morphology [48]. Criteria for scoring a germ cell as in the gonad included migration to abdominal segment 5 and presence within the correct

bilateral region to be incorporated into the gonad. Gonadal sheath cells were used to delimit the gonad boundaries when possible.

Genomic Sequence Analysis

Genomic DNA was extracted individually from seven *sctt* and five wild-type flies using a buffer containing 10mM Tris-HCl, 1mM EDTA, 25mM NaCl, and 10 mg/ml Proteinase K [53]. The *tre1* gene was PCR amplified using TripleMaster Taq DNA Polymerase (Eppendorf). PCR products were sequenced on both strands using an Applied Biosystems 3730xl DNA Analyzer (Iowa State University DNA Sequencing Facility).

Reverse Transcriptase PCR

0-8 hour embryos were collected, dechorionated using 50% bleach, homogenized in Trizol (Invitrogen), and total RNA isolated. DNA was removed from the total RNA using Turbo DNase (Ambion). First strand cDNA synthesis was performed using RETROscript First Strand cDNA Synthesis Kit for RT-PCR (Ambion). PCR was performed on the cDNA template using Taq DNA polymerase (Eppendorf). Lack of genomic contamination in the PCR amplification was confirmed using primers that spanned multiple exons, and the PCR products were sequenced.

Protein Secondary Structure Predictions

The Tre1 amino acid sequence of both wild type and *sctt* mutants were compared using the secondary structure prediction programs: SOSUI [30], TopPred [29], and TMHMM [31].

Engineering of the Amino Acid Substitution Cassettes

The $T^{+}G^{+}$ vector containing a 10 kb genomic fragment coding for both *tre1* and *Gr5a* was used in the generation of the amino acid-substituted constructs [27]. A 1700 base pair fragment containing the target sequence for nucleotide replacement was excised by digesting with SphI and StuI restriction endonucleases (NEB) and cloned into a modified pSP72 vector containing an inserted StuI restriction site and lacking a PstI site. The resulting pSP72 vector containing the 1700bp *tre1* insert was subsequently digested using PstI and

Bpu10I (NEB) to excise a 160 base pair fragment of *tre1* genomic DNA that housed the target sequence. Phosphorylated oligonucleotides were designed to reconstitute the 160 base pair fragment and insert an AloI restriction site directly into the target sequence region. Due to the positions of the codons in relation to the intron-exon border, two independent cassette vectors were designed with the AloI site engineered in different locations to allow the substitution of nucleotides encoding all 8 amino acids of interest. Vector sequencing confirmed the presence of the AloI restriction endonuclease sites in the reconstituted 1700 base pair subclone.

Creation of Constructs with Amino Acid Substitutions in Tre1

The pSP72 vector cassettes containing the target region with the engineered AloI restriction sites were digested with AloI (Fermentas). The AloI digest removed the AloI restriction site as well as 7 base pairs 5' and 12-13 base pairs 3' of the restriction enzyme recognition site. A pair of complementary phosphorylated oligonucleotides was designed to contain nucleotide changes to alter the amino acid sequence in the target region. The oligonucleotides were hybridized and ligated into the AloI-digested cassette. Sequencing confirmed the presence of the desired nucleotide substitutions. The pSP72 vector was then digested with SphI and StuI to excise the 1700 base pair fragment for insertion into the digested T^+G^+ vector to reconstitute the 10 kb genomic clone. The splice junctions and the amino acid substituted regions were sequenced to confirm the correct reading frame and construct composition.

Generation of Transgenic Flies and Fly Crosses

The engineered T^+G^+ constructs were injected into a w^{1118} host strain using a modification of the standard transformation protocol [54], as outlined by Nicholas Gompel. P{π25.7 D2-3 wc} was used a transposase source [55]. Transgenic stocks were established and the insertion sites of the transgenes within the genome were determined by inverse PCR [56]. The transgenic chromosomes with the amino acid substitutions were then crossed into the $tre1^{sctt}$ mutant background and stable homozygous or balanced heterozygous stocks established. To test for maternal rescue of germ cell migration, females homozygous for $tre1^{sctt}$ and carrying one or two copies of the transgenes were crossed to $tre1^{sctt}$ males and the offspring were assayed for germ cell migration defects.

Results

The sctt mutation disrupts normal germ cell migration

In *sctt* mutants, the migration of the germ cells to the gonads is severely disrupted [22]. Germ cells fail to migrate to and coalesce with somatic gonad precursor cells in embryos produced from a cross between a *sctt*/*sctt* female and a *sctt*/Y male (Figure 2.1B). Few, if any, of the germ cells reach the gonads (Table 2.1). The *sctt* allele is X-linked, recessive, and shows a maternal effect. One maternal copy of $sctt^{+}$ is sufficient to completely rescue germ cell migration (Figure 2.1D). In embryos from a homozygous *sctt* mutant female, germ cell migration can be rescued with a paternally supplied wild-type copy of the *sctt* gene [22]. Embryos derived from a cross between *sctt*/*sctt* females and a $sctt^{+}$ male display two phenotypes depending upon whether they inherit the wild-type X chromosome or the Y chromosome from the paternal genome. The *sctt* maternal^{-}/zygotic^{+} embryos are rescued for germ cell migration (Figure 2.1C), while those embryos with the *sctt* maternal^{-}/zygotic^{-} background have a severe germ cell migration phenotype (Figure 2.1B).

To establish a baseline for future transgenic rescue experiments and to better define the scattershot phenotype, germ cell counts were performed to determine the number of germ cells that reached the gonads in various *sctt* mutant backgrounds. Germ cell counts at stages 15-16 revealed an average of 14.7 germ cells reached the gonads in wild-type embryos (Table 2.1). These numbers are in agreement with previously published results from other genetic backgrounds [23-26]. While the total number of germ cells in *sctt* maternal^{-}/zygotic^{-} embryos is within the range of wild type, on average less than 1.0 germ cell per embryo reached the gonads. Rescue of germ cell migration by a paternally supplied copy of $sctt^{+}$ was complete, an average of 13.0 germ cells was observed in the gonads of *sctt* maternal^{-}/zygotic^{+} embryos, compared to 14.7 germ cells in the gonads for wild type. The *sctt* maternal^{-}/zygotic^{-} embryos from the same cross averaged 0.3 germ cells in the gonads (Table 2.1). This data supports and extends previous findings that germ cell migration to the gonads in *sctt* mutants is severely hindered but is successfully rescued when embryos express a wild-type copy of *sctt* [22].

The sctt mutation is an allele of tre1 that alters the splice acceptor site in intron 4 of tre1

The *sctt* maternal^{-}/zygotic^{-} germ cell phenotype suggested that the molecular defect causing this phenotype represents a severe loss of function. However, the molecular lesion causing the *sctt* mutation was unknown. It was known that *sctt* was an X-linked mutation [22]. To determine the location of the *sctt* gene, the *sctt* mutant chromosome was tested for complementation by crossing it to the Bloomington series of X chromosome deletion stocks. The *sctt* chromosome complemented all available deletions. Recombination mapping of *sctt* using a *w cv wy f* mapping X chromosome placed *sctt* within 1 map unit and distal to the *crossveinless* locus at 5A13. This suggested that *sctt* might be located in the 5A4 to 5A8-9 region, the gap between Df(1)JC70 (4C11;5A4) and Df(1)C149 (5A8-9;5C5). This mapping was consistent with the observation that the translocation Dp(1;Y)dx^{+}5 carrying genomic sequence from 4C11;6D8 of the X chromosome on the Y rescued the *sctt* cell migration phenotype. However, the Dp(1;Y)dx^{+}1 translocation of genomic sequence from 5A8-9;6D8 failed to rescue the *sctt* defect. The proximal limit of this region, 5A8-9, as defined by polytene chromosome breakpoints was in rough agreement with the recombination mapping, although the recombination mapping suggested that the gene might be in a slightly more proximal location. As noted below, the *sctt* gene corresponded to CG3171 located at 5A11-12.

Concurrent studies by Kunwar et al. revealed that the *sctt* chromosome failed to complement a deletion allele of *tre1*, ΔEP5 [27, 19]. However, since no molecular lesion had been identified, it remained unclear whether this represented allelic or non-allelic non-complementation. While the genetic mapping data suggested it was possible that *sctt* was an allele of *tre1,* genomic sequencing by Kunwar et al. found no evidence of a molecular lesion in the *tre1* coding region of the *sctt* mutant [19]. In addition, it was determined that *tre1* mRNA levels were not significantly decreased in the *sctt* background [28]. This argued that the *sctt* lesion was not resulting in a large change in *tre1* transcription. To identify if the *sctt* mutation was in *tre1*, the coding region along with introns of *tre1* was sequenced. Single adult *sctt* and wild type control flies were harvested and used for genomic templates. Primers directed at exons 2-7 of *tre1* were used in PCR to amplify and subsequently sequence this region. Nearly 2000 base pairs were sequenced on both DNA strands and a single base pair substitution was observed between *sctt* embryos and the wild type controls: an adenine within intron 4 was mutated to a thymine (Figure 2.2A).

The single base pair change of an adenine to a thymine within intron 4 could change the preferred splice acceptor site from AG to TG leading to improper splicing of the *tre1* RNA product. To test this, RNA was isolated from 0-8 hour old *sctt* or wild-type embryos and reverse transcriptase PCR was performed. Sequencing of the cDNA product confirmed that splicing of the wild type control occurred as predicted (Figure 2.2A). With the *sctt* template, intron/exon junctions through exon 4 were correctly spliced. However, directly following the altered splice acceptor site, 24 base pairs were missing from exon 5. The next suitable splice AG acceptor site in the sequence was used. This change results in an in-frame deletion of 8 amino acids, RYILIACH, from the protein (Figure 2.2B). The deleted amino acids are at the junction of the third transmembrane domain and the second intracellular loop of the Tre1 GPCR. Secondary structure analysis programs predict that the deletion shortens the second intracellular loop, while the remainder of the protein secondary structure is unaffected [29-31]. The removal of these eight amino acids from Tre1 has a significant impact on the function of the protein. This deletion includes two residues, R and Y, of the highly conserved E/N/DRY motif of rhodopsin family GPCRs. We concluded that *sctt* is an allele of *tre1*, and we will subsequently refer to it as $tre1^{sctt}$.

The arginine deleted in $Tre1^{sctt}$ is critical for primordial germ cell migration

The severe loss of function phenotype seen in $tre1^{sctt}$ mutants offered a unique opportunity to perform a detailed structure-function analysis of this region of the Tre1 GPCR. Amino acid sequence comparisons of this 8 amino acid region revealed conserved residues between Tre1 and human GPCRs known to function in cell migration (Figure 2.2C). Because this region is conserved, it was hypothesized that some or all of the amino acids deleted in $tre1^{sctt}$ were critical for Tre1 function. To test the hypothesis that specific amino acids in this region were necessary for Tre1 function, transgenic constructs were created where individual or small groups of amino acids in the deleted region were replaced with alanine. These amino acid substituted constructs were engineered into the T^+G^+ construct that contains 10 kb of Drosophila genomic sequence including the sequence for *tre1* and an adjacent gene, *Gr5a* [27]. Transgenic lines bearing transgenes designed to express altered forms of Tre1 were established. These transgenic chromosomes were then crossed into a $tre1^{sctt}/tre1^{sctt}$ genetic background. The function of the resulting Tre1 protein was assayed for maternal rescue of germ cell migration

by assaying germ cell migration in embryos from a *tre1*^{sctt}/*tre1*^{sctt}; P[transgene] mothers that were crossed to *tre1*^{sctt}/Y males.

To verify that maternally expressed transgenes can effectively rescue germ cell migration, the T^+G^+ construct was tested as a positive control [27]. In embryos from *tre1*^{sctt}/*tre1*^{sctt}; T^+G^+ mothers, germ cells successfully reached the gonads (Figure 2.3). An average of 23.0 germ cells were observed in the gonads, compared to an average of 0.3 germ cells detected in the gonads of embryos from *tre1*^{sctt}/*tre1*^{sctt} females lacking any transgene (Figure 2.4, Table 2.2). These values are statistically different ($P<0.0001$, Student's t-test). As a negative control, a transgene that reconstructed the *tre1*^{sctt} amino acid deletion was tested. A construct lacking all 8 amino acids was unable to rescue the *tre1*^{sctt} defect (Figure 2.3). A wild-type number of germ cells were counted within the embryo; however, an average of only 1.4 germ cells reached the gonads, similar to the no transgene control ($P>0.05$, Student's t-test) (Table 2.2).

To identify the critical amino acids within the RYILIACH deletion, constructs were designed with combinations of the original amino acids and alanine substitutions. If the Tre1 protein is dependent on any of these specific amino acids for proper germ cell migration, their replacement with alanine should disrupt Tre1 function and these constructs should fail to rescue the *tre1*^{sctt} phenotype. Alternatively, substitution of non-critical amino acids should not disrupt Tre1 function and the *tre1*^{sctt} germ cell migration phenotype should be rescued.

Germ cells successfully migrated to the gonads in embryos from *tre1*^{sctt}/*tre1*^{sctt} mothers carrying the RYAAAAAA, RYAAAACH, and RAILIACH constructs (Figures 2.3 and 2.4, Table 2.2). The average number of germ cells in the gonads was 20.9 and 20.4 for the two RYAAAAAA constructs, 16.2 for the RYAAAACH construct, and 23.0 for the RAILIACH construct (Table 2.2). This suggests that the seven amino acids, YILIACH, are not critical for Tre1 function in germ cell migration.

When the arginine was replaced with an alanine (AYILIACH), Tre1 function was not restored (Figure 2.3). An average of 0.4 and 1.1 germ cells were observed in the gonads of two different transgenic lines (Figure 2.4, Table 2.2). The phenotype appears similar to the *tre1*^{sctt} mutants lacking any transgene and the *tre1*^{sctt} reconstruction that fails to rescue the defect (Figure 2.3). This demonstrates that the arginine is essential for Tre1 function. While a shortening of the second intracellular loop may also contribute to the loss-of-function

observed with the *tre1*^{sctt} allele, the observation that restoring the length of the loop but replacing the arginine with an alanine results in a scattershot phenotype demonstrates a clear role for the arginine of the highly conserved E/N/DRY motif in proper germ cell migration in Drosophila.

Discussion

We report a genetic and functional analysis of Tre1, a GPCR that is essential for the migration of primordial germ cells in *Drosophila melanogaster*. Previous studies showed that the *sctt* chromosome failed to complement a chromosome carrying a deletion of *tre1* [19]. However, no associated molecular lesion had been discovered. Genomic and RT-PCR sequencing results presented in this study provide direct evidence that *sctt* is an allele of *tre1*. The *tre1*^{sctt} mutation results in an AG to TG mutation at a splice acceptor site that abolishes correct splicing resulting in the utilization a cryptic splice acceptor site and the of loss of the first 24 base pairs of exon 5 (Figure 2.2A). Eight amino acids are missing from the third transmembrane/second intracellular loop junction of the Tre1 GPCR while the rest of the protein proceeds in frame. Secondary structure modeling programs predict a shortening of the second intracellular loop with the remainder of the secondary structure unaffected.

The deleted region, RYILIACH, was systematically tested for restoration of Tre1 function using an alanine scan approach. Through the design of a modular cassette vector containing the *tre1* genomic sequence, the region encoding these 8 amino acids was manipulated to insert amino acid substitutions of each amino acid. Transgenic flies created using these constructs were assayed for their ability to maternally rescue the *tre1*^{sctt} germ cell migration defect in the context of developing embryos. Through this replacement approach, it was discovered that the 6 amino acids following the RY were dispensable for primordial germ cell migration. The RY<u>AAAAAA</u> and RY<u>AAA</u>ACH constructs restored germ cell migration back to wild-type levels.

The E/N/DRY motif is a hallmark of rhodopsin family GPCRs that has been studied extensively in tissue culture systems [1, 9]. As shown in Figure 2.3, an alanine replacement of the tyrosine in this domain fully rescued the *tre1*^{sctt} germ cell migration defect. It is possible that this highly conserved residue has important roles in other Tre1 functions, but it appears to be dispensable in early germ cell development. However, the neighboring arginine in Tre1 has critical roles in primordial germ cell migration. Two separate lines carrying

constructs that replaced the arginine with alanine were unable to rescue the *tre1*^{sctt} defect. Studies of GPCRs in other systems inform our interpretation of these results. Scheer et al. and Zhu et al. have examined the effects of analogous mutations in mammalian α1b-Adrenergic Receptors (R143A) and the m1 (R123N) and m2 (R121N) Muscarinic Receptors in tissue culture cells. They found that these mutations did not alter protein expression, but did impair GPCR signaling [8, 32]. We have not assayed Tre1 protein expression from these transgenes as epitope tags impaired Tre1 function (data not shown). However, other alanine substitutions within the region deleted in Tre1^{sctt} (RAILIACH, RYAAAACH, and RYAAAAAA) did provide Tre1 function. We therefore conclude that the failure of the arginine-substituted construct to rescue *tre1*^{sctt} loss of function is due to a lack of protein function rather than reduced expression.

The arginine of the E/N/DRY motif located towards the cytoplasmic side of the third transmembrane domain is conserved in 96% of rhodopsin family GPCRs [6, 7]. A sequence alignment of seven human rhodopsin family GPCRs involved in cell migration illustrates the invariant nature of this residue (Figure 2.2C). Rhodopsin family receptors are a set of highly diverse GPCRs both in their ligand binding ability and their elicited cellular responses. Cell culture experiments using nonconservative mutations of this residue result in defective signal transduction [8, 9, 32, 33]. In addition, others have suggested that this arginine can directly bind G proteins [34]. Alternatively, this arginine is proposed to be involved in the stability of receptor conformation [35-41]. While much research into the effects of amino acid alterations of the E/N/DRY motif has been performed in cell culture systems, this conserved motif has not been studied in the more complex and native signaling environment of an intact organism. We demonstrate that changing the arginine of this motif severely impacts the ability of this GPCR to function in primordial germ cell migration during Drosophila embryogenesis.

GPCR function has been identified as having critical roles in the directed migration of a variety of cell types. An emerging theme in cell migration is that ligands such as chemokines and phospholipids function as attractants for cells to specific locations. These ligands activate GPCRs on the receiving cell's surface to initiate a migratory response toward higher levels of the agonist. The phosopholipid sphingosine-1-phosphate receptors $S1P_{1-4}$ have been implicated in lymphocyte recirculation and tissue homing critical in adaptive immune responses [42]. Additionally, the SDF-1/CXCR4 ligand-GPCR pair has emerged as a conserved mechanism

regulating a variety of cell migrations in cancer, immune response, and in development. In breast cancer, it has been found that secondary site tumor colonization during cancer metastasis is not random, but rather there is directed migration of CXCR4-expressing cancer cells towards the SDF-1 ligand at common secondary sites such as lungs and bone marrow [43]. This common ligand-receptor pair has also been found to play a role in immune response and leukocyte trafficking to tumors [11]. Similar to Tre1, the CXCR4 receptor has roles in mouse, chick, and zebrafish germ cell migration [12-17, 44, 45]**.**

Tre1 is related to GPCRs with roles in the migration of cells (Figure 2.2C). While previous studies have used tissue culture systems to assess GPCR structure-function relationships, detailed molecular analyses of mutated receptors have not been performed in complex and dynamic systems where the cells under study are in contact with neighboring cells and responding to endogenous extracellular signals. This study provides conclusive evidence that the arginine of the highly conserved E/N/DRY motif is critical for Tre1 GPCR-mediated germ cell migration within the context of a developing organism. Given the conserved nature of this arginine residue in other GPCRs and its identification as a critical residue from cell culture studies [8, 9, 32, 33], it is likely that this arginine is critical to the function of many other GPCRs in a wide variety of cell types and organisms.

Acknowledgements

We thank the laboratories of John Carlson, Ken Howard, and Janice Fischer as well as the Bloomington Stock Center for fly stocks and reagents. Jo Anne Powell-Coffman, Elizabeth Asque, and Taryn Rowley offered helpful suggestions for improving the manuscript.

References

1. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol 63: 1256-1272.
2. Maudsley S, Martin B, Luttrell LM (2005) The origins of diversity and specificity in g protein-coupled receptor signaling. J Pharmacol Exp Ther 314: 485-494.
3. Drews J (2000) Drug discovery: a historical perspective. Science 287: 1960-1964.
4. Brink CB, Harvey BH, Bodenstein J, Venter DP, Oliver DW (2004) Recent advances in drug action and therapeutics: relevance of novel concepts in G-protein-coupled receptor and signal transduction pharmacology. Br J Clin Pharmacol 57: 373-387.

5. Marsolais D, Rosen H (2009) Chemical modulators of sphingosine-1-phosphate receptors as barrier-oriented therapeutic molecules. Nat Rev Drug Discov 8: 297-307.
6. Mirzadegan T, Benko G, Filipek S, Palczewski K (2003) Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. Biochemistry 42: 2759-2767.
7. Rosenkilde MM, Kledal TN, Schwartz TW (2005) High constitutive activity of a virus-encoded seven transmembrane receptor in the absence of the conserved DRY motif (Asp-Arg-Tyr) in transmembrane helix 3. Mol Pharmacol 68: 11-19.
8. Zhu SZ, Wang SZ, Hu J, el-Fakahany EE (1994) An arginine residue conserved in most G protein-coupled receptors is essential for the function of the m1 muscarinic receptor. Mol Pharmacol 45: 517-523.
9. Scheer A, Costa T, Fanelli F, De Benedetti PG, Mhaouty-Kodja S et al. (2000) Mutational analysis of the highly conserved arginine within the Glu/Asp-Arg-Tyr motif of the alpha(1b)-adrenergic receptor: effects on receptor isomerization and activation. Mol Pharmacol 57: 219-231.
10. Berchiche YA, Chow KY, Lagane B, Leduc M, Percherancier Y et al. (2007) Direct assessment of CXCR4 mutant conformations reveals complex link between receptor structure and G(alpha)(i) activation. J Biol Chem 282: 5111-5115.
11. Koizumi K, Hojo S, Akashi T, Yasumoto K, Saiki I (2007) Chemokine receptors in cancer metastasis and cancer cell-derived chemokines in host immune response. Cancer Sci 98: 1652-1658.
12. Richardson BE, Lehmann R Mechanisms guiding primordial germ cell migration: strategies from different organisms. Nat Rev Mol Cell Biol 11: 37-49.
13. Ara T, Nakamura Y, Egawa T, Sugiyama T, Abe K et al. (2003) Impaired colonization of the gonads by primordial germ cells in mice lacking a chemokine, stromal cell-derived factor-1 (SDF-1). Proc Natl Acad Sci U S A 100: 5319-5323.
14. Molyneaux KA, Zinszner H, Kunwar PS, Schaible K, Stebler J et al. (2003) The chemokine SDF1/CXCL12 and its receptor CXCR4 regulate mouse germ cell migration and survival. Development 130: 4279-4286.
15. Dumstrei K, Mennecke R, Raz E (2004) Signaling pathways controlling primordial germ cell migration in zebrafish. J Cell Sci 117: 4787-4795.
16. Stebler J, Spieler D, Slanchev K, Molyneaux KA, Richter U et al. (2004) Primordial germ cell migration in the chick and mouse embryo: the role of the chemokine SDF-1/CXCL12. Dev Biol 272: 351-361.
17. Raz E, Mahabaleshwar H (2009) Chemokine signaling in embryonic cell migration: a fisheye view. Development 136: 1223-1229.
18. Coffman CR (2003) Cell migration and programmed cell death of Drosophila germ cells. Ann N Y Acad Sci 995: 117-126.
19. Kunwar PS, Starz-Gaiano M, Bainton RJ, Heberlein U, Lehmann R (2003) Tre1, a G protein-coupled receptor, directs transepithelial migration of Drosophila germ cells. PLoS Biol 1: E80.
20. Kamps AR, Coffman CR (2005) G protein-coupled receptor roles in cell migration and cell death decisions. Ann N Y Acad Sci 1049: 17-23.
21. Kunwar PS, Sano H, Renault AD, Barbosa V, Fuse N et al. (2008) Tre1 GPCR initiates germ cell transepithelial migration by regulating Drosophila melanogaster E-cadherin. J Cell Biol 183: 1-12.
22. Coffman CR, Strohm RC, Oakley FD, Yamada Y, Przychodzin D et al. (2002) Identification of X-linked genes required for migration and programmed cell death of Drosophila melanogaster germ cells. Genetics 162: 273-284.
23. Sonnenblick BP (1941) Germ Cell Movements and Sex Differentiation of the Gonads in the Drosophila Embryo. Proc Natl Acad Sci U S A 27: 484-489.
24. Sonnenblick BP (1950) The early embryology of Drosophila melanogaster. New York: John Wiley.
25. Underwood EM, Caulton JH, Allis CD, Mahowald AP (1980) Developmental fate of pole cells in Drosophila melanogaster. Dev Biol 77: 303-314.
26. Hay B, Ackerman L, Barbel S, Jan LY, Jan YN (1988) Identification of a component of Drosophila polar granules. Development 103: 625-640.
27. Dahanukar A, Foster K, van der Goes van Naters WM, Carlson JR (2001) A Gr receptor is required for response to the sugar trehalose in taste neurons of Drosophila. Nat Neurosci 4: 1182-1186.
28. Burnett C (2005) The role of lipid phosphate phosphatases and their putative receptors in germ cell migration and survival in Drosophila melanogaster. PhD Thesis University College London.

29. von Heijne G (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol 225: 487-494.
30. Hirokawa T, Boon-Chieng S, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14: 378-379.
31. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580.
32. Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. Embo J 15: 3566-3578.
33. Jones PG, Curtis CA, Hulme EC (1995) The function of a highly-conserved arginine residue in activation of the muscarinic M1 receptor. Eur J Pharmacol 288: 251-257.
34. Acharya S, Karnik SS (1996) Modulation of GDP release from transducin by the conserved Glu134-Arg135 sequence in rhodopsin. J Biol Chem 271: 25406-25411.
35. Ballesteros J, Kitanovic S, Guarnieri F, Davies P, Fromme BJ et al. (1998) Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. J Biol Chem 273: 10445-10453.
36. Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L et al. (2001) Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. J Biol Chem 276: 29171-29177.
37. Angelova K, Fanelli F, Puett D (2002) A model for constitutive lutropin receptor activation based on molecular simulation and engineered mutations in transmembrane helices 6 and 7. J Biol Chem 277: 32202-32213.
38. Greasley PJ, Fanelli F, Rossier O, Abuin L, Cotecchia S (2002) Mutagenesis and modelling of the alpha(1b)-adrenergic receptor highlight the role of the helix 3/helix 6 interface in receptor activation. Mol Pharmacol 61: 1025-1032.
39. Shapiro DA, Kristiansen K, Weiner DM, Kroeze WK, Roth BL (2002) Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. J Biol Chem 277: 11441-11449.
40. Flanagan CA (2005) A GPCR that is not "DRY". Mol Pharmacol 68: 1-3.
41. Zhang M, Mizrachi D, Fanelli F, Segaloff DL (2005) The formation of a salt bridge between helices 3 and 6 is responsible for the constitutive activity and lack of hormone responsiveness of the naturally occurring L457R mutation of the human lutropin receptor. J Biol Chem 280: 26169-26176.
42. Matloubian M, Lo CG, Cinamon G, Lesneski MJ, Xu Y et al. (2004) Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. Nature 427: 355-360.
43. Muller A, Homey B, Soto H, Ge N, Catron D et al. (2001) Involvement of chemokine receptors in breast cancer metastasis. Nature 410: 50-56.
44. Doitsidou M, Reichman-Fried M, Stebler J, Koprunner M, Dorries J et al. (2002) Guidance of primordial germ cell migration by the chemokine SDF-1. Cell 111: 647-659.
45. Knaut H, Werz C, Geisler R, Nusslein-Volhard C (2003) A zebrafish homologue of the chemokine receptor Cxcr4 is a germ-cell guidance receptor. Nature 421: 279-282.
46. Ueno K, Ohta M, Morita H, Mikuni Y, Nakajima S et al. (2001) Trehalose sensitivity in Drosophila correlates with mutations in and expression of the gustatory receptor gene Gr5a. Curr Biol 11: 1451-1455.
47. Fischer-Vize JA, Rubin GM, Lehmann R (1992) The fat facets gene is required for Drosophila eye and embryo development. Development 116: 985-1000.
48. Campos-Ortega J, Hartenstein V (1997) The Embryonic Development of Drosophila melanogaster. New York: Springer-Verlag.
49. Johansen KM, Johansen J (2004) Studying nuclear organization in embryos using antibody tools. Methods Mol Biol 247: 215-234.
50. Simon JA, Sutton CA, Lobell RB, Glaser RL, Lis JT (1985) Determinants of heat shock-induced chromosome puffing. Cell 40: 805-817.
51. Holmes AL, Raper RN, Heilig JS (1998) Genetic analysis of Drosophila larval optic nerve development. Genetics 148: 1189-1201.

52. Yamada Y (2007) Deciphering molecular mechanisms that regulate programmed cell death of primordial germ cells in Drosophila melanogaster. PhD Thesis Iowa State University.
53. Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW et al. (1993) Type I repressors of P element mobility. Genetics 135: 81-95.
54. Spradling AC, Rubin GM (1982) Transposition of cloned P elements into Drosophila germ line chromosomes. Science 218: 341-347.
55. Karess RE, Rubin GM (1984) Analysis of P transposable element functions in Drosophila. Cell 38: 135-146.
56. Bellen HJ, Levis RW, Liao G, He Y, Carlson JW et al. (2004) The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics 167: 761-781.

CHAPTER 3. FUNCTIONS OF THE TRE1 G PROTEIN-COUPLED RECEPTOR AND THE WUNEN AND WUNEN2 LIPID PHOSPHATE PHOSPHATASES IN PRIMORDIAL GERM CELL MIGRATION

Introduction

As precursors to eggs and sperm, the primordial germ cells are the vital genetic link to the next generation. The movements of these cells during development have been extensively studied in *Drosophila melanogaster* and have proven to be useful models for studies of cell migration and programmed cell death. In Drosophila, the primordial germ cells are the first cells to cellularize and are formed at the posterior pole of the embryo [1]. During gastrulation, the primordial germ cells are moved into the developing posterior midgut primordium in close association with the underlying epithelium. The primordial germ cells begin active migration at stage 9 by dispersing into individual cells and migrating across the epithelium of the posterior midgut [2-4]. Once the primordial germ cells have exited the posterior midgut, they separate into two clusters of cells in the mesoderm and move to and align with the somatic gonadal precursor cells. By stage 14, the primordial germ cells have coalesced with the somatic gonadal precursor cells to form the embryonic gonads [4,5].

One of the first proteins involved in Drosophila primordial germ cell migration is the G protein-coupled receptor (GPCR) Trapped in endoderm-1 (Tre1), an orphan receptor of the Rhodopsin family of GPCRs that functions within the primordial germ cells [4,6-9]. Using a loss-of-function allele of *tre1, tre1*$^{\Delta EP5}$, it was shown that *tre1* is necessary for the primordial germ cells to properly exit the posterior midgut [4,9], being required for the polarization, dispersion and transepithelial migration of the primordial germ cells. In *tre1*$^{\Delta EP5}$ mutants, the primordial germ cells fail to polarize, do not disperse into individual cells and do not migrate across the posterior midgut epithelium. The result is that very few primordial germ cells successfully migrate to the developing gonads [4]. A second loss-of-function allele of *tre1*, *scattershot* (*tre1*^{sctt}), also shows a primordial germ cell migration defect. In *tre1*^{sctt}, only one germ cell on average migrates properly to the developing gonads, while the rest remain in ectopic locations [8].

In addition to Tre1, the lipid phosphate phosphatases Wunen (Wun) and Wunen2 (Wun2) are involved in

primordial germ cell development with roles in the transepithelial migration of the primordial germ cells, their bilateral segregation and their survival [10-14]. Although it is known that Tre1, Wun and Wun2 have important roles at multiple points in Drosophila primordial germ cell development, the mechanisms of action and the signaling networks that regulate primordial germ cell migration and survival are poorly understood. The ligand(s) for Tre1 are not known, and GPCRs are an extremely large and diverse family of proteins that have numerous functions including cell migration, survival, proliferation and metastasis [7,15,16]. In addition, the substrates that Wun and Wun2 dephosphorylate are not known. Interestingly, lipid phosphate phosphatases regulate the activities of important lipid signaling molecules such as lysophosphatidic acid, sphingosine 1-phosphate, phosphatidic acid, and ceramide-1-phosphate by catalyzing the removal of phosphates. Removal of phosphates alters the signaling properties of these lipids that can be ligands for or modulators of GPCRs [17-21].

The goal of this study was to obtain more information about the signaling networks involved in Drosophila primordial germ cell migration. Using both genetic and yeast two-hybrid screens, a possible connection between the Tre1 GPCR and the lipid phosphate phosphatases Wun and Wun2 was established. Late stage germ cell phenotypes of zygotic loss-of-function mutant embryos for *wun* and *wun2* appear similar to germ cell phenotypes of $tre1^{sctt}$ mutant embryos. In the alleles studied, primordial germ cell migration was severely disrupted with most primordial germ cells left ectopic to the gonads. The yeast two-hybrid screen revealed that Tre1 interacted with both Wun and Wun2 when Tre1 was the bait. The results presented here, in addition to the previously published roles of Tre1, Wun and Wun2 in primordial germ cell migration, suggest that Tre1, Wun and Wun2 may be part of an integrated signaling network governing primordial germ cell migration in Drosophila.

Materials and Methods

Fly Stocks

The $tre1^{sctt}$ allele was generated in an EMS mutagenesis screen [22]. The w^{1118}, P {w^{+}, *fat facets-lacZ*} (faf^{X}) line was used as the control as it was the parental strain used in the EMS mutagenesis screen that yielded $tre1^{sctt}$ [23]. The following lines from the Bloomington Stock Center were used in the loss-of-function screen: $CG7047^{KG05089}$, $crk^{KG00336}$, $gprk2^{06936}$, $gprk2^{EY09213}$, $kurtz^{c01503}$, $neur^{KG06174}$, $neur^{11}$, $neur^{A101}$, $pxt^{EY03052}$,

src64B^{KG00213}, *src64B*^{EY16432}, *tao-1*^{EP1455}, *trio*^{6A}, *trio*^{KG06642}, *wun*^{GL}, *wun*^{CE}, and *wun*^{k10201}. The *jafrac2*^{f01922} and *tao-1*^{e04532} lines were obtained from the Harvard Exelixis Collection. Ruth Lehmann kindly provided the Df(2R)*NYX-D15*/CyO, P[*ry*$^{+}$*, ftz-lacZ*] and *wun2*^{EP2650ex34}/CyO, P[*ry*$^{+}$*, ftz-lacZ*] stocks.

Immunohistochemistry and Germ Cell Counts

Embryos were dechorionated in 50% bleach, fixed in 4% paraformaldehyde, and devitellinized using a 1:1 mixture of heptane:methanol [24]. The primary antibodies were chicken anti-Vasa (1:10,000, a gift from Ken Howard and Camilla Burnett)[25], rabbit anti-Vasa (1:10,000, a gift from Ruth Lehmann), mouse anti-β-Galactosidase (1:1,000, Promega), mouse anti-EYA (1:25, 10H6 Developmental Studies Hybridoma Bank)[26], and mouse anti-Hindsight (1:10, 1G9 Developmental Studies Hybridoma Bank)[27]. The secondary antibodies were biotinylated anti-chicken IgG (1:500, Vector Laboratories), biotinylated anti-rabbit IgG (1:500, Vector Laboratories) and biotinylated anti-mouse IgG (1:500, Vector Laboratories). The ABC Elite Kit from Vector Laboratories was used to complex avidin: horseradish peroxidase to the secondary antibodies, and 3,3'-diaminobenzidine tetrahydrochloride (DAB) was used to detect peroxidase activity. The Vector VIP Substrate Kit for Peroxidase (VIP) was also used as a substrate to visualize the peroxidase activity. For germ cell counts at stages 10-11 both VIP and DAB substrates were used. The primordial germ cells were labeled using chicken anti-Vasa and the VIP substrate on the first day. The anterior and posterior midguts were labeled using anti-Hindsight and the DAB substrate on the second day. When scoring stages 10-11 embryos, the primordial germ cells were placed into two categories: primordial germ cells that were crossing the posterior midgut and primordial germ cells that had exited the posterior midgut. A primordial germ cell was considered crossing the posterior midgut if it was on the same focal plane as Hindsight-stained midgut cells and at the periphery of the stained posterior midgut. Germ cell counts were performed on stages 5-6, stages 10-11 and stages 15-16 embryos using differential interference contrast microscopy. Embryos were staged according to published criteria [5].

Split-Ubiquitin Yeast Two-Hybrid

Protein-protein interactions were examined using the split-ubiquitin yeast two-hybrid system, according to

the manufacturer's instructions (DUALsystems, Schlieren, Switzerland) [28-31]. The split-ubiquitin yeast two-hybrid system allows the detection of interactions between integral membrane proteins. In split-ubiquitin yeast two-hybrid assays, an interaction between two hybrid proteins (bait and prey) leads to the reconstitution of an ubiquitin molecule at the plasma membrane. Cleavage by ubiquitin-specific proteases releases the LexA-VP16 transcription factor from a membrane anchor allowing the activation of *HIS3*, *ADE2*, and *lacZ* reporter genes. The strength of the interaction was quantified by co-transforming yeast with plasmids encoding both hybrid proteins and determining the fraction of transformants with sufficient reporter gene activity to allow growth into colonies on selective medium. Growth observed using a negative control plasmid (pDL2-*ALG5*) was considered background growth. Colony counts were normalized by subtracting the number of colonies observed with pDL2-*ALG5* between the three independent experiments. Results were compared using a Student's t-test. Transformants were assessed qualitatively for β-galactosidase activity by their ability to develop a blue color on media containing 240 mg/L of 5-bromo-4-chloro-indolyl-galactopyranoside. Additionally, interaction was confirmed by growing double transformants on histidine- and adenine-free medium supplemented with 5 mM of 3-amino-1,2,4-triazole (aminotriazole, AT).

Constructs when Tre1 was the bait

Bait vectors were constructed by cloning the full-length cDNA of *tre1* into the vectors pBT3-*STE* and pBT3-*SUC* to yield plasmids pBT3-*STE*-*tre1* and pBT3-*SUC*-*tre1*, respectively. Functionality of bait constructs was examined by co-transforming *S. cerevisiae* strain NMY51 with bait plasmids and positive (pAI-*ALG5*) and negative (pDL2-*ALG5*) control prey vectors. Prey vectors were constructed by cloning the full-length cDNAs of *wun* and *wun2* into the vector pPR3-N to yield plasmids pPR3-N-*wun* and pPR3-N-*wun2*. Transforming *S. cerevisiae* strains with the non-interacting *PEX11* bait construct with both pPR3-N-*wun* and pPR3-N-*wun2* prey vectors tested specificity of interactions. No interaction between Wun or Wun2 preys and Pex11 control bait could be observed, confirming the specificity of the assay. Fourteen additional prey vectors were constructed by cloning full-length cDNAs into pPR3-C or pPR3-N vectors (Table 3.5). All bait and prey vectors were sequenced to confirm correct composition and reading frame using an Applied Biosystems 3730xl DNA Analyzer (Iowa State University DNA Sequencing Facility). cDNAs were obtained through the Drosophila Genomics Resource Center [32,33].

Constructs when Wunen and Wunen2 were the baits

Bait vectors were constructed by cloning the full-length cDNAs of *wun* and *wun2* into the vectors pBT3-N, pBT3-*STE*, pBT3-*SUC* and pBT3-C to yield eight total bait vectors. Prey vectors were constructed by cloning the full-length cDNA of *tre1* into the vectors pPR3-*STE* and pPR3-*SUC* to yield plasmids pPR3-*STE*-*tre1* and pPR3-*SUC*-*tre1*, respectively. These constructs were sequenced to confirm correct construct composition and reading frame.

Results and Discussion

A candidate gene loss-of-function genetic screen identifies wunen and wunen2 mutants as displaying germ cell phenotypes similar to tre1^{sctt}

The Tre1 GPCR is an orphan receptor that plays a central role in Drosophila primordial germ cell migration. Despite extensive efforts to elucidate the mechanisms of Tre1 signaling, components of this signaling network, both upstream and downstream of this receptor, have remained elusive and connections between the known elements fragmentary [4,9]. To identify additional genes involved in Tre1-mediated signaling, as well as discover new genes involved in Drosophila primordial germ cell migration, a screen was performed that used loss-of-function alleles of genes with potential roles in primordial germ cell migration. Selection of candidate genes was based on literature searches, expression of the genes in primordial germ cells [34,35], involvement in GPCR-mediated signaling and/or cell migration in other systems.

Mutant alleles of eleven candidate genes were examined for primordial germ cell phenotypes using immunohistochemistry with a chicken anti-Vasa antibody [25] to label the primordial germ cells of embryos at stages 15-16, a point after the completion of primordial germ cell migration to the gonads (Table 3.1). Of these eleven genes, only *wun* and *wun2* loss-of-function alleles disrupted germ cell development. Other candidate genes either gave a normal germ cell phenotype or the disruption in germ cell development was associated with the genetic background and not the candidate gene (data not shown). The *wun* and *wun2* genes were of particular interest since they gave germ cell phenotypes that were similar to *tre1^{sctt}* at stages 15-16 (Figure 3.1).

Five different alleles of *wun* and *wun2* were studied in this loss-of-function screen: *wun^{GL}*, *wun^{CE}*,

Df(2R)*NYX-D15*, wun^{k10201}, and $wun2^{EP2650ex34}$. The wun^{GL}, wun^{CE} and Df(2R)*NYX-D15* alleles are deletions in both *wun* and *wun2* [10,12,14,36]. The wun^{k10201} allele is a hypomorphic allele of *wun* [37], and $wun2^{EP2650ex34}$ is a *wun2* RNA-null allele [11,12]. Additionally, transheterozygote embryos between wun^{GL} and Df(2R)*NYX-D15* were constructed since wun^{GL}/Df(2R)*NYX-D15* has been reported as the strongest allelic combination of *wun* and *wun2* [13,14]. For subsequent studies, Df(2R)*NYX-D15*, wun^{GL} and wun^{GL}/Df(2R)*NYX-D15* were used and will be referred to as *wun wun2* mutants.

At stages 15-16 Df(2R)*NYX-D15*/Df(2R)*NYX-D15,* wun^{GL}/wun^{GL}, and wun^{GL}/Df(2R)*NYX-D15* embryos have similar primordial germ cell phenotypes to $tre1^{sctt}$ embryos (Figure 3.1). To further compare these phenotypes, germ cell counts were performed to analyze the location and number of the primordial germ cells at stages 15-16 (Table 3.2). In all four *wun wun2* genetic backgrounds, most of the primordial germ cells remain in positions ectopic to the gonads with few primordial germ cells migrating to the gonads. In $tre1^{sctt}$ embryos only 1.0 primordial germ cell, on average, migrates properly to the gonads while an average of 18.6 primordial germ cells are ectopic to the gonads. This is similar to Df(2R)*NYX-D15*/Df(2R)*NYX-D15,* wun^{GL}/wun^{GL} and the two wun^{GL}/Df(2R)*NYX-D15* lines where 2.2, 4.7, 4.6 and 3.2 primordial germ cells on average migrate properly to the gonads, and 34.7, 26.2, 29.4, and 29.7 primordial germ cells, on average, are left in ectopic positions.

The similarity of the phenotypes of the $tre1^{sctt}$ and *wun wun2* alleles is intriguing and suggests that the roles these genes play in primordial germ cell development may be linked. It is known that maternal *tre1*, *wun* and *wun2* are all involved in the initial phases of primordial germ cell migration [7,8,10,11,13,14,36,37]. *tre1* is involved during the first active phase of primordial germ cell migration and is required for the proper polarization, dispersal and transepithelial migration of the primordial germ cells [4,9]. *wun* and *wun2* are functionally redundant in both their somatic and germline functions [10,11,13,14,36,37]. In the primordial germ cells *wun* and *wun2*, like *tre1*, function early in the migration of these cells across the posterior midgut epithelium. Germline *wun* and *wun2* are also involved in germ cell-germ cell repulsion that is thought to help the primordial germ cells disperse, cross the midgut and separate into two clusters of cells in the mesoderm [11]. In somatic tissues, *wun* and *wun2* are expressed in the central nervous system (CNS), the epidermis and on the ventral side of the midgut, all regions of the embryo that wild-type primordial germ cells avoid. In these regions, somatic *wun* and *wun2* help establish a repulsive environment, which restricts the primordial germ cells

to an area of the embryo around the developing gonads [11,13,14,36,37].

The analyses presented here demonstrate the similarity in stages 15-16 germ cell phenotypes between the *tre1*^{sctt} and *wun wun2* alleles. This suggests that the roles Tre1, Wun and Wun2 have in dispersal and transepithelial migration of primordial germ cells may overlap and that these proteins may function together as part of a common signaling network.

The initiation of primordial germ cell migration is delayed in tre1*^{sctt} *embryos but not in wun wun2 embryos

tre1^{sctt} and *wun wun2* embryos analyzed at stages 15-16 have intriguingly similar primordial germ cell phenotypes. At this point in development, wild-type primordial germ cells have coalesced with the somatic gonadal precursor cells to form the embryonic gonads. To address the question of whether *tre1*, *wun* and *wun2* are involved in connected signaling networks, the hypothesis that the primordial germ cell phenotypes will be similar at earlier stages, when the primordial germ cells are beginning active migration at stage 10, was tested.

To determine if *tre1*^{sctt} and *wun wun2* embryos have similar primordial germ cell phenotypes at stages 10-11, the location of the primordial germ cells was analyzed and counted in control (*faf*^{X}), *tre1*^{sctt} and *wun wun2* embryos. Embryos in this experiment were double antibody stained with chicken anti-Vasa to label the primordial germ cells and anti-Hindsight to label the nuclei of the anterior and posterior midgut, allowing precise determination of the position of the primordial germ cells in relation to the midgut epithelium (Figure 3.2) [27]. In Drosophila, primordial germ cells migrate as individual cells [4] and do not all migrate out of the posterior midgut at the same time. To assess the dynamics of primordial germ cell migration out of the posterior midgut, germ cell counts were performed at two time points. The first time point was early stage 10, when the primordial germ cells in wild-type embryos have already initiated migration and have begun to cross the epithelium of the posterior midgut. The second time point was late stage 10 – early stage 11. At this point, most if not all of the primordial germ cells in wild-type embryos have exited the posterior midgut and have entered the overlying mesoderm. At each of the two time points, the primordial germ cells were counted and placed into two categories, primordial germ cells crossing the posterior midgut and primordial germ cells that have exited the posterior midgut. The primordial germ cells that were still in the posterior midgut were not scored, as they are tightly associated with each other and difficult to accurately assay. This analysis was

performed on control (*faf*^{X}), *tre1*^{sctt} and *wun wun2* (Df(2R)*NYX-D15*/Df(2R)*NYX-D15*) embryos (Table 3.3). As expected, primordial germ cells in control embryos have initiated active migration at early stage 10 and most have completed migrating out of the posterior midgut by late stage 10 – early stage 11. At early stage 10, on average 5.6 primordial germ cells are crossing the posterior midgut and 16.8 primordial germ cells have exited the posterior midgut. At late stage 10 – early stage 11, an average of 0.1 primordial germ cells are crossing the posterior midgut and 34.6 primordial germ cells have exited the posterior midgut. In control embryos, all primordial germ cells were scored at late stage 10 – early stage 11 as none remained in the posterior midgut. Interestingly, germ cell counts on *tre1*^{sctt} embryos reveal the exit of the primordial germ cells from the posterior midgut is delayed. On average, 2.1 primordial germ cells are crossing the posterior midgut and 1.9 primordial germ cells have exited the posterior midgut at early stage 10 in *tre1*^{sctt} embryos. By late stage 10 – early stage 11, an average of 3.8 primordial germ cells are crossing the posterior midgut and 5.1 primordial germ cells have exited the posterior midgut. At this stage, there are still primordial germ cells in the posterior midgut of *tre1*^{sctt} embryos.

From this data, two things are concluded: first, primordial germ cell exit from the posterior midgut is delayed in *tre1*^{sctt} embryos; second, at least some of the primordial germ cells are capable of exiting the posterior midgut in *tre1*^{sctt} embryos. During stage 11 the posterior midgut cells undergo an epithelial to mesenchymal transition and the posterior midgut architecture collapses [5], thus it is not possible to reliably score whether the primordial germ cells that remain in the posterior midgut in *tre1*^{sctt} embryos at early stage 11 ever actively cross the midgut epithelium later in development (data not shown).

At stages 10-11, germ cell counts have revealed that the primordial germ cells in *wun wun2* embryos exit the midgut similar to control embryos (Table 3.3). At early stage 10, an average of 5.4 primordial germ cells are crossing the posterior midgut and 12.8 primordial germ cells have exited the posterior midgut in *wun wun2* embryos. By late stage 10 – early stage 11, on average there are 7.1 primordial germ cells crossing the midgut and 28.4 primordial germ cells that have exited the posterior midgut. Although all of the primordial germ cells were counted at late stage 10 – early stage 11 in control embryos, some of the primordial germ cells were still in the posterior midgut at this time in *wun wun2* embryos and were not scored. This could be due to the lack of *wun* and *wun2* expression on the ventral side of the midgut at this time in development. The lack of *wun* and

wun2 expression on the ventral side of the midgut allows the primordial germ cells in a *wun wun2* embryo to exit the midgut in all directions. This phenotype made it difficult to resolve if either some primordial germ cells were crossing the posterior midgut or if the primordial germ cells were within the posterior midgut. Thus, the number of primordial germ cells scored as crossing the posterior midgut epithelium in *wun wun2* mutants may be under-reported. The results from the germ cell counts at stages 10-11 in *wun wun2* embryos are consistent with previous observations where the primordial germ cells in *wun wun2* embryos exit the posterior midgut [13,14,36].

The primordial germ cells in *wun wun2* mutant embryos, like seen in the control embryos, initiate migration out of the posterior midgut at the correct time during development. However, the primordial germ cells in $tre1^{sctt}$ mutant embryos are delayed in their exit from the posterior midgut. These results suggest that the origin of the similar germ cell phenotypes at stages 15-16 for *wun wun2* and $tre1^{sctt}$ mutant embryos is not the same. Reviewing the germ cell counts at stages 15-16 revealed that *wun wun2* mutant embryos have approximately twice as many primordial germ cells as control and $tre1^{sctt}$ embryos. To determine if this was due to differences in the genetic backgrounds, germ cell counts at stages 5-6 were performed.

$tre1^{sctt}$ and wun wun2 mutant embryos have similar numbers of primordial germ cells at stages 5-6

The primordial germ cells cellularize at stage 4 at the posterior pole of the developing Drosophila embryo. During stage 4, the primordial germ cells undergo 0-2 cell divisions, resulting in approximately 40 primordial germ cells at stage 5 [1,5]. As the primordial germ cells in wild type Drosophila have stopped dividing at this point in development and gastrulation begins at stage 6, germ cell counts at stages 5-6 are an excellent time to determine the number of germ cells initially formed in the embryo. Control (faf^{X}), $tre1^{sctt}$, Df(2R)*NYX-D15*/Df(2R)*NYX-D15* and wun^{GL}/wun^{GL} embryos were used in this study and the primordial germ cells were stained with a rabbit anti-Vasa antibody. All genetic backgrounds have similar numbers of germ cells formed at stages 5-6 (Table 3.4). On average, 35.3 primordial germ cells are present in faf^{X}, 34.5 in $tre1^{sctt}$, 36.5 in Df(2R)*NYX-D15*/Df(2R)*NYX-D15* and 37.7 in wun^{GL}/wun^{GL}. These results show that faf^{X}, $tre1^{sctt}$, and *wun wun2* embryos begin with nearly the same number of primordial germ cells. In addition, these results suggest that the similar germ cell phenotypes of *wun wun2* and $tre1^{sctt}$ mutant embryos at stage 15-16 likely

arise by different mechanisms.

Although the germ cell counts at stages 5-6 suggest *faf*^{X}, *tre1*^{sctt}, and *wun wun2* embryos begin with the same number of primordial germ cells, it is possible that the primordial germ cells continue to divide in *wun wun2* embryos as gastrulation begins. However, the germ cell counts at stages 10-11 (Table 3.3) and stages 15-16 (Table 3.2) argue this is not the case. Additionally, Sano et al. have reported an average of 37.0 primordial germ cells in *wun*^{GL}/Df(2R)*NYX-D15* embryos at stage 10 [13]. This is in agreement with the data presented here and suggests the primordial germ cells in a *wun wun2* embryo do not continue to divide after stage 6.

Tre1 can interact with both Wunen and Wunen2 in a split-ubiquitin yeast two-hybrid assay

Concurrent to the genetic loss-of-function screen, a yeast two-hybrid screen was performed to identify additional proteins that may be components in the Tre1-signaling network. Due to the transmembrane nature of Tre1, the split-ubiquitin yeast two-hybrid system was used (DUALsystems, Schlieren, Switzerland) [28-31]. Candidate proteins were chosen based on literature, those tested in the loss-of-function screen, and genes that have been shown to be involved in primordial germ cell migration. Sixteen candidate genes were cloned into the prey vectors and tested for an interaction with Tre1. Two of these candidates, Wun and Wun2, showed a significant interaction with Tre1 (Table 3.5).

Using two different Tre1 baits, Tre1 interacts with both Wun and Wun2 as measured by growth on adenine- and histidine-free medium supplemented with aminotriazole (5 mM), as well as by the induction of the β-Galactosidase reporter (Figure 3.3). To quantitate the growth, the fraction of transformants that grew into colonies on selective media (-Ade-His) was determined (Figure 3.4). For the Wun prey, the percentage of colonies that grew was 13.6 ± 2.1% (average ± standard error of the mean) and 12.7 ± 1.6% for the pBT3-Ste-Tre1 and pBT3-Suc-Tre1 baits, respectively. Similar results were observed with the Wun2 prey where 10.1 ± 0.3% and 9.0 ± 3.8% colonies grew with pBT3-Ste-Tre1 and pBT3-Suc-Tre1 baits, respectively. The interaction seen is significantly different than the negative control prey with either Tre1 bait ($P<0.05$, Student's t-test).

To further test these results, the baits and the preys were switched. Wun and Wun2 were constructed as baits and Tre1 was constructed as a prey. Eight different baits were constructed and, at this time, only two have

been tested, pBT3-N-Wun and pBT3-N-Wun2. These baits showed self-activation, so in order to test for an interaction with the Tre1 prey, aminotriazole (5 mM for pBT3-N-Wun and 1 mM for pBT3-N-Wun2) need to be added to the media to suppress the background growth. When inhibiting the self-activation with aminotriazole, an interaction between the Wun and Wun2 baits and the Tre1 prey could not be detected (data not shown). The self-activating nature of pBT3-N-Wun and pBT3-N-Wun2 was problematic. Once the additional baits are analyzed, the ability of Tre1 to interact with Wun and Wun2 can be assessed.

If the other Wun and Wun2 baits do not interact with Tre1, it would suggest that the yeast two-hybrid results with Tre1 as a bait were false positives. A lack of interaction with the Wun and Wun2 baits could also mean that the interaction between Tre1, Wun and Wun2 may be transient and not easily seen in a yeast two-hybrid assay. Other experiments would need to be performed to determine if this is the case and potential experiments and alternative hypotheses will be discussed in Chapter 5.

If the other Wun and Wun2 baits show an interaction with Tre1, the results would suggest, taken together with the established roles of Tre1, Wun and Wun2 in the initiation of primordial germ cell migration, that Tre1, Wun and Wun2 may act in a closely coupled signaling network during the early phases of primordial germ cell migration in Drosophila. Tre1, Wun and Wun2 appear to play similar roles in the dispersal of the primordial germ cells prior to their exit from the posterior midgut. In $tre1^{\Delta EP5}$ mutant embryos, which lack maternal and zygotic *tre1*, the primordial germ cells do not disperse into individual cells and do not exit the posterior midgut [4]. Although the dispersal of the primordial germ cells is necessary, it is not sufficient to cause transepithelial migration. In embryos lacking maternal DE-cadherin, the primordial germ cells disperse early, however the primordial germ cells are delayed in their transepithelial migration across the posterior midgut. Furthermore, removing *tre1* function ($tre1^{\Delta EP5}$) in a maternal loss-of-function DE-cadherin mutant, the primordial germ cells showed precocious dispersal, but failed to exit the midgut [4]. Comparably, in embryos lacking maternal and zygotic *wun* and *wun2*, the primordial germ cells do not disperse into individual cells and do not cross the midgut epithelium [11]. Similar to what was seen in $tre1^{\Delta EP5}$ mutants, the dispersal of these *wun wun2* mutant primordial germ cells (caused by a loss of maternal DE-cadherin) was not sufficient to cause transepithelial migration [11]. Given these observations in addition to the yeast-two hybrid data, it is possible that the Tre1 GPCR and the lipid phosphate phosphatases Wun/Wun2 mediate the initiation of primordial germ cell

migration and/or the sensing of guidance cues by either closely coupled signaling pathways or by a direct modulation of Tre1 signaling by Wun and/or Wun2.

References

1. Sonnenblick BP (1941) Germ cell movements and sex differentiation of the gonads in the Drosophila embryo. PNAS 27: 484-489.
2. Callaini G, Giovanna M, Riparbelli, Dallai R (1995) Pole cell migration through the gut wall of the Drosophila embryo - analysis of cell interactions. Developmental Biology 170: 365-375.
3. Jaglarz M, Howard K (1995) The active migration of Drosophila PGC. Development 121: 3495-3503.
4. Kunwar PS, Sano H, Renault AD, Barbosa V, Fuse N, et al. (2008) Tre1 GPCR initiates germ cell transepithelial migration by regulating Drosophila melanogaster E-cadherin. J Cell Biol 183: 1-12.
5. Campos-Ortega J, Hartenstein V (1997) The Embryonic Development of *Drosophila melanogaste*r. New York: Springer-Verlag. 405 p.
6. Coffman CR (2003) Cell migration and programmed cell death of Drosophila germ cells. Annals New York Academy of Sciences 995: 117-126.
7. Kamps AR, Coffman C (2005) G Protein-Coupled Receptor Roles in Cell Migration and Cell Death Decisions. Annals of the New York Academy of Sciences 1049: 17-23.
8. Kamps AR, Pruitt MM, Herriges JC, Coffman CR (2010) An Evolutionarily Conserved Arginine is Essential for Tre1 G Protein-Coupled Receptor Function During Germ Cell Migration in Drosophila melanogaster. PLoS ONE 5: e11839.
9. Kunwar P, Starz-Gaiano M, Bainton R, Heberlein U, Lehmann R (2003) Tre1, a GPCR, directs transepithelial migration of Drosophila germ cells. Plos Biol 1: 372-384.
10. Hanyu-Nakamura K, Kobayashi S, Nakamura A (2004) Germ cell-autonomous Wunen2 is required for germline development in Drosophila embryos. Development 131: 4545-4553.
11. Renault AD, Kunwar PS, Lehmann R (2010) Lipid phosphate phosphatase activity regulates dispersal and bilateral sorting of embryonic germ cells in Drosophila. Development 137: 1815-1823.
12. Renault AD, Sigal Y, Morris A, Lehmann R (2004) Soma-Germ Line Competition for Lipid Phosphate Uptake Regulates Germ Cell Migration and Survival. SCIENCE 305: 1963-1966.
13. Sano H, Renault A, Lehmann R (2005) Control of lateral migration and germ cell elimination by the Drosophila melanogaster lipid phosphate phosphatases Wunen and Wunen 2. The Journal of Cell Biology 171: 675-683.
14. Starz-Gaiano M, Cho N, Forbes A, Lehmann R (2001) Spatially restricted activity of a Drosophila lipid phosphatase guides migrating germ cells. Development 128: 983-991.
15. Koizumi K, Hojo S, Akashi T, Yasumoto K, Saiki I (2007) Chemokine receptors in cancer metastasis and cancer cell-derived chemokines in host immune response. Cancer Science 98: 1652-1658.
16. Radeff-Huang J, Seasholtz T, Matteo R, Brown J (2004) G Protein Mediated Signaling Pathways in Lysophospholipid Induced Cell Proliferation and Survival. Journal of Cellular Biochemistry 92: 949-966.
17. Pyne S, Kong K, Darroch P (2004) Lysophosphatidic acid and sphingosine 1-phosphate biology: the role of lipid phosphate phosphatases. Seminars in Cell & Developmental Biology 15: 491-501.
18. Pyne S, Lee SC, Long J, Pyne NJ (2009) Role of sphingosine kinases and lipid phosphate phosphatases in regulating spatial sphingosine 1-phosphate signalling in health and disease. Cellular Signalling 21: 14-21.
19. Samadi N, Bekele R, Capatos D, Venkatraman G, Sariahmetoglu M, et al. (2011) Regulation of lysophosphatidate signaling by autotaxin and lipid phosphate phosphatases with respect to tumor progression, angiogenesis, metastasis and chemo-resistance. Biochimie 93: 61-70.
20. Sciorra VA, Morris AJ (2002) Roles for lipid phosphate phosphatases in regulation of cellular signaling. Biochim Biophys Acta 1582: 45-51.
21. Spiegel S, Milstien S (2003) Sphingosine-1-phosphate: an enigmatic signalling lipid. Nature Reviews Molecular Cell Biology 4: 397-407.

22. Coffman CR, Strohm R, Oakley F, Yamada Y, Przychodzin D, et al. (2002) Identification of X-linked genes required for migration and programmed cell death of Drosophila melanogaster germ cells. Genetics 162: 273-284.
23. Fischer-Vize J, Rubin G, Lehmann R (1992) The fat facets gene is required for Drosophila eye and embryo development. Development 116: 985-1000.
24. Johansen K, Johansen J (2004) Studying nuclear organization in embryos using antibody tools. Methods in Molecular Biology 247: 215-234.
25. Burnett C (2005) The role of lipid phosphate phosphatases and their putative receptors in germ cell migration and survival in Drosophila melanogaster. PhD Thesis University College London.
26. Bonini NM, Leiserson WM, Benzer S (1993) The eyes absent gene: genetic control of cell survival and differentiation in the developing Drosophila eye. Cell 72: 379-395.
27. Yip ML, Lamka ML, Lipshitz HD (1997) Control of germ-band retraction in Drosophila by the zinc-finger protein HINDSIGHT. Development 124: 2129-2141.
28. Johnsson N, Varshavsky A (1994) Split ubiquitin as a sensor of protein interactions in vivo. Proceedings of the National Academy of Sciences of the United States of America 91: 10340-10344.
29. Stagljar I, Korostensky C, Johnsson N, te Heesen S (1998) A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. Proc Natl Acad Sci U S A 95: 5187-5192.
30. Thaminy S, Auerbach D, Arnoldo A, Stagljar I (2003) Identification of Novel ErbB3-Interacting Factors Using the Split-Ubiquitin Membrane Yeast Two-Hybrid System. Genome Research 13: 1744-1753.
31. Iyer K, Burkle L, Auerbach D, Thaminy S, Dinkel M, et al. (2005) Utilizing the Split-Ubiquitin Membrane Yeast Two-Hybrid System to Identify Protein-Protein Interactions of Integral Membrane Proteins. SCIENCE STKE 275: PI3.
32. Stapleton M, Liao G, Brokstein P, Hong L, Carninci P, et al. (2002) The Drosophila Gene Collection: Identification of Putative Full-Length cDNAs for 70% of D. melanogaster Genes. Genome Research 12: 1294-1300.
33. Hoskins RA, Landolin JM, Brown JB, Sandler JE, Takahashi H, et al. (2011) Genome-wide analysis of promoter architecture in Drosophila melanogaster. Genome Research 21: 182-192.
34. Tomancak P, Beaton A, Weiszmann R, Kwan E, Shu S, et al. (2002) Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol 3: RESEARCH0088.
35. Tomancak P, Berman B, Beaton A, Weiszmann R, Kwan E, et al. (2007) Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biol 8: R145.
36. Zhang N, Zhang J, Cheng Y, Howard K (1996) Identification and genetic analysis of wunen, a gene guiding Drosophila melanogaster germ cell migration. Genetics 143: 1231-1241.
37. Zhang N, Zhang J, Purcell K, Cheng Y, Howard K (1997) The Drosophila protein Wunen repels migrating germ cells. Nature 385: 64-67.

CHAPTER 4: A COMPUTATIONAL STUDY ON THE TRE1 G PROTEIN-COUPLED RECEPTOR

Introduction

G protein-coupled receptors (GPCRs) are the largest class of membrane proteins, accounting for 2% of genes in the human genome [1-3]. In general, GPCRs are responsible for modulating signals from the extracellular space and transducing these stimuli into intracellular signaling cascades. GPCRs are involved in processes including cell movement, neurotransmission and olfaction, and can also be involved in disease progression with roles in metastasis, angiogenesis, cell proliferation and inflammation. Since GPCRs are involved in maintaining homeostasis as well as disease progression, GPCRs are an important group of proteins to study.

Structurally, all GPCRs share the feature of having seven transmembrane α-helices (TM1-TM7) connected by three intracellular and three extracellular loops. Phylogenetic trees of the TM helices were used to classify GPCRs into five families: Rhodopsin, Secretin, Glutamate, Adhesion and Frizzled/Taste2 [1,2].

Despite the importance of GPCRs, there are not many GPCR crystal structures available. Obtaining crystal structures is critical for a more complete understanding of the structure-function relationships within the GPCR superfamily. The limited number of crystal structures available is not from a lack of effort. GPCRs are inherently difficult to crystallize due to the transmembrane nature of GPCRs and the fact that GPCRs are typically expressed in low levels within cells. GPCRs, like other transmembrane proteins, require a membrane-like environment to remain in a properly folded conformation. The required presence of a membrane makes the overexpression and subsequent purification of GPCRs challenging. Crystallization of GPCRs in lipidic mesophases (also known as the liquid crystalline phase) is a more promising technique than the typical crystallization in detergent solutions due to the amphiphilic nature of GPCRs [4,5]. Lipidic mesophases have been used to successfully crystallize a number of GPCRs [6-9].

The first GPCR crystal structure, bovine rhodopsin, was determined in 2000 [10], with nearly seven years passing before a crystal structure for another GPCR was published. There are currently 13 inactive, non-opsin GPCRs crystallized, all representing the Rhodopsin family of GPCRs: human β_2-adrenergic receptor [6,11], turkey β_1-adrenergic receptor [12], human A_{2A}-adenosine receptor [8], human chemokine (CXC motif)

receptor 4 (CXCR4) [9], human D_3 dopamine receptor [7], human H_1 histamine receptor [13], human M2 muscarinic receptor [14], rat M3 muscarinic receptor [15], army worm $S1P_1$ receptor [16], human κ-opiod receptor [17], mouse μ-opiod receptor [18], mouse δ-opioid receptor [19] and human nociception/orphanin FQ receptor [20]. Additionally, some of these proteins (Rhodopsin, Adenosine and β Adrenergic) have been crystallized in activated, ligand-bound states [21-26]. Although there has been a breakthrough in the ability to crystallize GPCRs, more structures are needed due to the diversity in structure and function within the five families of GPCRs. A broader set of GPCR structures will enable investigators to learn more about how the structure of each type of GPCR relates to their specific function(s).

Due to the difficulties of GPCR crystallization, protein structure prediction programs and molecular dynamics (MD) simulations are frequently used to investigate the structures of GPCRs. However, the prediction of three-dimensional structures for GPCRs is still a challenge despite the numerous available prediction programs. There are currently three computational techniques available to generate a three-dimensional structural prediction of a protein: homology modeling, threading and *de novo* or *ab initio* modeling. Homology modeling builds a three-dimensional structure by first identifying an evolutionarily related homologous protein with a known structure to use as a template. The program then aligns the amino acid sequence of the protein of interest to the amino acid sequence of the chosen template and finally builds the model [27-29]. The relatively low number of GPCR crystal structures is a major limitation to homology modeling. A lack of diverse structures means that a majority of GPCRs will still lack a homologous protein to use as a template. When the template protein and the protein sequence of interest share 50% or more sequence identity it is possible to build a highly accurate model [27,28]. However, when the sequence identity is below 30%, the protein structure built will likely more closely resemble the template structure than the native structure of the protein [30]. The sequence identity between crystallized GPCRs and other known GPCRs is often below 30% [31]. Due to the prevalence of low sequence identity, it is suggested that both sequence identity and structural information be used when choosing the template protein to use [31]. Threading, similar to homology modeling, is a template-based approach to modeling. The first step in threading is to search for evolutionary relatives to the protein sequence of interest. This is commonly accomplished with Position-Specific Iterative Basic Local Alignment Search Tool (PSI-BLAST) [32]. PSI-BLAST generates a sequence profile, which is

used by a secondary structure predictor, like PSIPRED [33], to determine the secondary structure of the protein sequence of interest. Both the secondary structure and the sequence profile from PSI-BLAST are used in a threading algorithm to identify template proteins from the Protein Data Bank that have similar protein folds to the sequence of interest. Templates used in threading may show no evolutionary relationship [29]. The use of multiple templates, creating a chimeric GPCR, has been shown to provide a more accurate model than using a single protein template [31,34,35]. Multiple templates can be used in both homology modeling and threading. *Ab initio* modeling builds a three-dimensional protein model from sequence information alone, without using a template structure, based upon the assumption that the protein structure will assume the lowest free energy conformation [27]. *Ab initio* modeling can work well for proteins with less than 120 amino acids [29]. Although there are three different ways to build a protein structure, some current modeling programs use a combination of approaches to generate structure predictions [29]. The accuracy of the final model is linked to the template(s) chosen and some approaches to generating a protein structure work better on certain proteins or parts of proteins than other approaches [36,37].

With only 14 distinct GPCR proteins crystallized, it can be challenging to find a suitable template(s) to use in the modeling software. Recently, three web servers have become available specific to modeling GPCRs: GPCR-Sequence Structure Feature Extractor (GPCR-SSFE) [38], GPCR-ModSim [39] and GPCR-Iterative Threading ASSEmbly Refinement (GPCR-ITASSER) [40-42]. GPCR-SSFE is a database in addition to a homology modeling program that creates homology models of GPCRs using multiple templates and the program MODELLER [38,43,44]. The ability to use multiple templates is significant since the use of multiple templates with MODELLER has been shown to give more accurate homology models than using one template [35]. However, GPCR-SSFE does not model the loop regions of GPCRs.

GPCR-ModSim is another server that also allows researchers to model GPCRs using MODELLER [39,43,44]. GPCR-ModSim is specific to GPCRs because the user has the option of choosing whether to align their GPCR sequence with seven inactive-like crystallized GPCRs or five active-like crystallized GPCRs. GPCR-ModSim aligns the sequence and shows the percent identity with the available templates. The user can then choose which template to use and GPCR-ModSim generates a homology model. Once a homology model

is generated, the user has the option of submitting it for MD simulations in a solvated 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipid bilayer [39].

GPCR-ITASSER is a web server that allows for protein structure prediction [40-42]. GPCR-ITASSER takes the initial GPCR sequence and identifies evolutionary relatives using PSI-BLAST and secondary structures using PSIPRED. The results from PSI-BLAST and PSIPRED are used by the Local Meta-Threading Server (LOMETS) to find potential templates in the Protein Data Bank. Any sequence without a matched template is modeled using an *ab initio* helix-modeling program. Additional restraints to the protein structure are incorporated through the use of the online database GPCRRD (GPCR Research Database), which contains experimental restraints from other GPCR databases and literature [40-42]. The *ab initio* modeling, results from threading and restraints from the GPCRRD are all used to assemble and build the structural model.

Trapped in endoderm-1 (Tre1) is a GPCR required for proper *Drosophila melanogaster* primordial germ cell migration [45-47]. In embryos carrying a severe partial loss-of-function allele of the *tre1* gene, $tre1^{sctt}$, the primordial germ cells scatter across the posterior half of the embryo rather than forming two gonads. The molecular lesion in $tre1^{sctt}$ RNA is a point mutation that results in an in-frame deletion of eight amino acids, RYILIACH, which are located at the junction of the third transmembrane domain and second intracellular loop of the Tre1 protein [47]. However, it is not known how the loss of these amino acids affects Tre1 protein structure. Two of the amino acids deleted in the $Tre1^{sctt}$ protein are part of the highly conserved D/ERY motif in the Rhodopsin family of GPCRs [3,48]. The arginine in the D/ERY motif is conserved in 96% of Rhodopsin family GPCRs [49] and can form a salt bridge with TM6 in numerous GPCRs [7,50-60]. Although the exact role of the salt bridge in GPCR stabilization is unclear [61], it is possible that the arginine in the Tre1 protein is responsible for forming a salt bridge to TM6 in wild type Tre1 ($Tre1^{+}$) protein.

In this study, the hypothesis that the arginine of the D/ERY motif is critical to maintain Tre1 protein structure was tested. As there is no crystal structure of Tre1 available, the GPCR-ModSim [39] and the GPCR-ITASSER [40-42] web servers were used to generate protein structure predictions of both $Tre1^{+}$ and $Tre1^{sctt}$ proteins. The NAMD simulation package [62] was used to perform MD simulations on both $Tre1^{+}$ and $Tre1^{sctt}$ embedded in a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) lipid bilayer. MD simulations were run for 200 ns for each of the four systems studied. At this time, the simulations are not well sampled and

have not converged and it is clear that each of the simulations must be run for longer times. Yet from the data collected thus far, it appears as if a salt bridge can form in $Tre1^{+}$ and, while it is possible for an alternative salt bridge to form in $Tre1^{sctt}$, it is not as stable as the salt bridge in $Tre1^{+}$.

Methods

Protein Structure Predictions

Protein structure predictions were generated for $Tre1^{+}$ and $Tre1^{sctt}$ using GPCR-ModSim [39] and GPCR-ITASSER [40-42]. Using squid rhodopsin (2Z73) as a template, ten homology models were generated by GPCR-ModSim. The best model (as judged by the lowest Discrete Optimized Protein Energy (DOPE-HR) score) was chosen for further refinement of the loop regions. The DOPE-HR score is used to assess the quality of the models generated. Five models were generated after loop refinement and again the best model was chosen based on DOPE-HR score. The GPCR-ModSim model generated for $Tre1^{+}$ will be referred to as mtre1 and the model generated for $Tre1^{sctt}$ will be referred to as msctt. Five models were generated by GPCR-ITASSER and the best model as judged by the confidence score (C-score) was chosen for this study. C-score is an estimation of the structural prediction and is based on the threading alignments from LOMETS and convergence during the structural refinements [29]. The GPCR-ITASSER model generated for $Tre1^{+}$ will be referred to as gtre1 and the model generated for $Tre1^{sctt}$ will be referred to as gsctt.

Building a System to Reflect a Drosophila Cellular Membrane Environment

The protein structure predictions (mtre1, msctt, gtre1 and gsctt) were embedded in a solvated (0.15 M NaCl) and pre-equilibrated POPE lipid bilayer using the Membrane Builder in the CHARMM-GUI [63,64]. The proteins were embedded in a POPE lipid bilayer since phosphoethanolamine is the major phospholipid in Drosophila cell membranes [65].

For the mtre1 system, 101 Na^{+} and 112 Cl^{-} ions were added to neutralize the system and the system contained 37,557 water molecules. The upper and lower leaflets of the membrane contained 141 and 137 POPE lipids, respectively. The mtre1 system had a total of 153,870 atoms. The msctt system contained 110 Na^{+} and 120 Cl^{-} ions, 40,839 water molecules, 141 and 137 POPE lipids on upper and lower leaflets of the membrane,

respectively, and 163,593 total atoms. The gtre1 system contained 69 Na^+ and 80 Cl^- ions, 26,139 waters molecules, 140 and 137 POPE lipids on upper and lower leaflets of the membrane, respectively, and 119,427 total atoms. The gsctt system contained 68 Na^+ and 78 Cl^- ions, 26,102 waters molecules, 139 and 140 POPE lipids on upper and lower leaflets of the membrane, respectively, and 119,423 total atoms.

Molecular Dynamics Simulations

MD simulations were performed using the NAMD 2.8 simulation package [62]. NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. The CHARMM22 [66,67] and CHARMM36 [68] force fields were used for protein and lipids, respectively, and water molecules were described using TIP3P [69]. All systems were simulated at 310 K. Temperature and pressure were held constant with Langevin dynamics [62] and the Nose-Hoover Langevin piston [70,71]. Particle-mesh Ewald was used to calculate electrostatic interactions [72] and a 12 Å cut-off for van der Waals interactions was used. Each system was simulated on three compute nodes, each containing one Intel(R) Xeon(R) X5650 CPU (6 cores at 2.67GHz), two Nvidia C2070 graphical processing units (GPUs) and 24 GB of RAM connected by QDR QLogic Infiniband.

After building the systems with the Membrane Builder in the CHARMM-GUI, six short (25 or 100 ps) equilibrium simulations were performed to gradually equilibrate the systems. Details for the equilibrium simulations can be found in [63]. Briefly, positional harmonic restraints were used on the protein backbone, protein side chains and ions. Additional harmonic restraints were used on the water molecules, to prevent water molecules from entering the hydrophobic region of the membrane, and the lipid head groups, to keep the lipid head groups level with the Z-axis. The restraints were reduced at each subsequent equilibrium simulation. The first two simulations used the NVT (constant volume and temperature) ensemble and the last four equilibrium simulations used the NPAT (constant pressure, area and temperature) ensemble. A timestep of 1 fs was used for the first three equilibrium simulations, which were 25 ps each. The last three equilibrium simulations used a 2 fs timestep and were run for 100 ps each [63]. Production runs began after the systems were equilibrated and

used an NPT (constant pressure and temperature) ensemble and a 2 fs timestep. Harmonic restraints were not used in the production runs. Each production simulation ran for approximately 200 ns.

Data Analysis

Visual Molecular Dynamics 1.9 (VMD) [73] was used to visualize the trajectories and to perform the all-to-all RMSD calculations, the salt bridge analysis and the solvent accessible surface area (SASA) calculations. A 1.4 Å radius probe was used in the SASA calculation. The LOOS (Lightweight Object-Oriented Structure) analysis library [74] was used to calculate the decorrelation time [75-77] and assess convergence [78]. The transmembrane Cα atoms were used in the decorrelation time and convergence calculations. The Voronoi Tesselation and Monte Carlo (VTMC) integration method was used to calculate the surface area per lipid in all model systems [79] to ensure the systems maintained a biologically relevant, fluid phase lipid bilayer.

Amino Acid Numbering

Amino acid residues are labeled either using the single-letter or three-letter code for the amino acid followed by the absolute sequence number. For example, arginine 134 is labeled R134 or ARG134. Figure 4.1 depicts the specific arginines used in this study and shows how the amino acids are numbered. $Tre1^{sctt}$ is missing eight amino acids compared to $Tre1^{+}$; however, the absolute sequence number of amino acid residues studied in this protein is still used. For example, R233 is in the same location as R225 (Figure 4.1 B).

Results and Discussion

Protein structure prediction

The Tre1 GPCR is most closely related to melatonin, histamine and serotonin receptors, based on sequence similarity [45]. Unlike most Rhodopsin family GPCRs, Tre1 does not contain the highly conserved D/ERY motif in TM3 and instead contains an NRY motif. Like Tre1, melatonin receptors have an NRY motif. However, the Tre1 protein sequence is unlike melatonin receptors and has the NPxxY motif in TM7 while melatonin receptors have the NAxxY motif in TM7. The NPxxY motif is another highly conserved sequence

motif in Rhodopsin family GPCRs that is thought to have roles in the activation of the receptor [48]. The NRY motif in Tre1 is of particular interest since two of the amino acids in this motif (RY) are missing in the mutant form of the protein, $Tre1^{sctt}$. The arginine of the D/E/NRY motif is the most conserved residue of the motif [49] and is critical for proper function of the Tre1 GPCR [47].

The amino acid sequences for $Tre1^{+}$ (NCBI accession number AAF46059) and $Tre1^{sctt}$ were used for protein structure predictions using GPCR-ModSim [39] and GPCR-ITASSER [40-42]. Both GPCR-ModSim and GPCR-ITASSER are web servers for GPCR protein structure prediction, however the web servers differ in the approach taken to generating a protein structure prediction. GPCR-ModSim automates using MODELLER to model GPCRs [39] while GPCR-ITASSER uses multiple threading programs as well as the GPCRRD to generate protein structure predictions [40-42]. Both web servers were used in this study to generate two independent protein structure predictions for both $Tre1^{+}$ and $Tre1^{sctt}$.

$Tre1^{+}$ and $Tre1^{sctt}$ were modeled to the seven inactive-like GPCRs on the GPCR-ModSim web server. From the multiple sequence alignment generated, it was clear that a few of the available GPCR crystal structures could be used as a template to model $Tre1^{+}$ and $Tre1^{sctt}$. However, GPCR-ModSim allows only one template to be chosen, and for both $Tre1^{+}$ and $Tre1^{sctt}$ the template chosen was squid rhodopsin. Squid rhodopsin showed the highest total sequence similarity to $Tre1^{+}$ and $Tre1^{sctt}$ (17.4% for $Tre1^{+}$ and 16.7% for $Tre1^{sctt}$) (Table 4.1). Using squid rhodopsin as a template, models for both $Tre1^{+}$ and $Tre1^{sctt}$ were generated using MODELLER. The models chosen for further study using MD simulations were chosen for the lowest DOPE-HR score and are named mtre1 ($Tre1^{+}$) and msctt ($Tre1^{sctt}$).

The second set of independent protein structure predictions for $Tre1^{+}$ and $Tre1^{sctt}$ were built using GPCR-ITASSER [40-42]. The amino acid sequences for $Tre1^{+}$ and $Tre1^{sctt}$ were submitted to GPCR-ITASSER and used in the local threading server to find template proteins. Both $Tre1^{+}$ and $Tre1^{sctt}$ were modeled to Substance P, human β_2-adrenergic receptor, bovine rhodopsin and human A_{2A} adenosine receptor. $Tre1^{+}$ was also modeled to turkey β_1-adrenergic receptor and β_2-adrenergic receptor-Gs protein complex, while $Tre1^{sctt}$ was also modeled to squid rhodopsin. The best models, based on C-score, were chosen for further study in MD simulations. The models generated by GPCR-ITASSER are named gtre1 ($Tre1^{+}$) and gsctt ($Tre1^{sctt}$).

The four models, mtre1, msctt, gtre1, and gsctt, chosen for further study are shown in Figure 4.2A. At first glance, all four of the models look similar, but there are distinct differences. Namely, the C-terminal tails in mtre1 and msctt (the red chains) are not structured while they are in gtre1 and gsctt. Additionally, helices 5 and 6 (yellow and gold chains) are roughly the same length as the other five helices in mtre1 and msctt. In gtre1 and gsctt, helices 5 and 6 are extended or longer than the other five helices. Also, intracellular loop 2 (green) has different structures in mtre1 and gtre1. In mtre1, this loop region is unstructured. In contrast, there is a short helix in intracellular loop 2 in gtre1. Intracellular loop 2 is of interest since it is the location of some of the residues missing in $Tre1^{sctt}$. The significance of these differences is not yet known.

As GPCRs exist in a cellular, membrane environment, the four different protein structure predictions were inserted into a solvated POPE lipid bilayer using the Membrane Builder in the CHARMM-GUI [63,64]. The final solvated, membrane systems are named the same as the structural predictions, mtre1, msctt, gtre1 and gsctt. Each system was subjected to 200 ns of MD and an example of the mtre1 system after MD is shown in Figure 4.3. The final protein structures after 200 ns of dynamics in a lipid bilayer are shown in Figure 4.2B (the lipids, water and ions are not shown). In general, the protein structures have not changed drastically and some of the differences between the initial structures generated by GPCR-ModSim and GPCR-ITASSER are still present. While the C-terminal tails have structure in all models, helices 5 and 6 remain different between the models generated by GPCR-ModSim and GPCR-ITASSER. Also, the short helix in intracellular loop 2 remains in gtre1 after 200 ns of dynamics. Another noticeable difference between the models generated by GPCR-ModSim and those generated by GPCR-ITASSER is TM2 (the light blue chain). During the course of the dynamics, TM2 changes from a single helix to a disconnected helix in both mtre1 and msctt. After 200 ns of dynamics, TM2 in mtre1 and msctt breaks into a short loop and then reforms a helix on the extracellular half of the helix. Again, while the significance of these differences remains to be understood, it is interesting that even after 200 ns of MD, some of the structural differences between the model systems were not resolved. This could mean that the models generated by the different modeling programs represent different protein conformations of Tre1.

It is important that GPCRs are simulated in systems that mimic the natural environment of GPCRs and allow for the systems to maintain a biologically relevant state. Force fields are used in molecular dynamics

simulations to describe the different molecules that comprise the environment of the system. Initially, an older force field, CHARMM27, was used to describe the POPE molecules. However, the descriptions of POPE molecules by this force field were not accurate and the bilayer transitioned to a gel phase-like surface area per lipid at the fluid phase temperature (data not shown). Experimental work has shown that at 310 K, the POPE molecules in a POPE lipid bilayer should maintain a surface area of 59.75 – 60.75 $Å^2$ [80]. To correct for the lower surface area per lipid, a newer force field, CHARMM36 [68], was used. CHARMM36 contains updates to the description of lipid molecules and with this force field POPE-only systems maintain the correct surface area per lipid (data not shown). To confirm the correct surface area per lipid were maintained in mtre1, msctt, gtre1 and gsctt, the Voronoi Tesselation and Monte Carlo (VTMC) integration method [79] was used (Table 4.2). This method allows for calculation of the surface area per lipid in membrane-lipid systems. VTMC calculates the surface area per boundary and non-boundary lipids. Non-boundary lipids are described as those lipids not interacting with atoms of the protein. It is important to make the distinction between lipid types (boundary versus non-boundary) since lipids interacting with atoms of the protein will have a decreased surface area per lipid. Results with VTMC confirmed that the non-boundary lipids in each of the model systems maintained the correct surface area per lipid, ranging from 59.8 – 61.0 $Å^2$ (Table 4.2).

Statistical analyses and convergence

Like any set of data, MD simulations are prone to statistical errors. The errors can be from inaccuracies in the model or inadequate sampling. For this reason, it is important to report the statistical uncertainty of values determined from simulations. In order to calculate the statistical uncertainty of different values in a simulation, the number of independent samples within a single simulation needs to be known. It has been suggested that estimation of a value of interest based on less than 20 statistically independent samples is considered unreliable [81]. To calculate the number of independent samples within a simulation, the decorrelation time must be calculated. The decorrelation time is the minimum amount of time that must pass between data points in the trajectory for the data points to become decorrelated, or for the data points to lose "memory" of the other data point.

There are many methods available to calculate the decorrelation time of a simulation. This study used the decorrelation time and effective sample size methods developed by the Zuckerman lab [76,77], as well as the block covariance overlap method (BCOM) from the Grossfield lab [78]. The effective sample size gives the degrees to which a simulation has sampled the conformational space of the protein and BCOM is a method used to measure the extent of convergence of a simulation.

The decorrelation times as estimated by the structural histogram analysis and by the automated effective sample size are shown in Table 4.3. The Cα atoms of the transmembrane regions of the proteins were the only atoms considered in these calculations. The approximate decorrelation times for the four model systems studied ranges from 11.7 – 141.2 ns. As the decorrelation times are estimates, the longest decorrelation time should be chosen as the estimated decorrelation time to use [78]. Based on these criteria, the decorrelation times for mtre1, msctt, gtre1 and gsctt are 119.5 – 141.2, 125.2 – 138.5, 105.6 – 132.9, and 123.6 – 140.1 ns, respectively. The effective sample size calculated for each system ranged from 1.0 – 1.2 (data not shown), indicating that the systems have insufficient sampling and the associated decorrelation times are not accurate.

To confirm the results from the tools developed by the Zuckerman lab, decorrelation times were also calculated using the bootstrapped block covariance overlap method (BBCOM) and BCOM for the transmembrane Cα atoms (Table 4.4). Three different decorrelation times are calculated in this method and are described as fast, medium and slow decorrelation times [78]. Similar to the decorrelation times calculated above, the slowest (longest) decorrelation time should be used for the approximate decorrelation time. The corresponding decorrelation times for mtre1, msctt, gtre1 and gsctt are 83.0, 75.8, 128.2 and 81.8 ns, respectively. However, the final ratio of BCOM/BBCOM indicates a lack of convergence. A final ratio of 1.0 indicates convergence and the final ratios for the systems studied ranged from 1.3 – 1.6. The BCOM/BBCOM ratios are determined from the plots of BCOM/BBCOM over time (data not shown).

The results from the structural histogram analysis, the automated effective sample size calculation and the blocked covariance overlap method indicate that the systems, at this point, have not converged. This means that statistics generated from the data will not be sufficient to draw conclusions. However, these results are not surprising. Microsecond simulations (or longer) with other GPCRs did not show convergence using these same methods [78]. Since the systems in this study have not converged, the values presented in the rest of this

chapter will either omit error bars or will be a more qualitative assessment of the simulations. Currently, the simulations have run for approximately 200 ns. The simulations are continuing to run and will continue running until statistical errors can be calculated.

Global movements of the model systems

Root mean squared deviation (RMSD) over the course of the simulation was computed for the Cα atoms of the transmembrane regions of the proteins (Figure 4.4). Plotting RMSD versus time is used to measure structural convergence. As RMSD depends on the protein and the system it interacts with, there is no set value to what RMSD should be and the RMSD value cannot be compared between simulation runs. However, in RMSD versus time plots it is common to see a rapid rise in RMSD due to the initial addition of heat to the simulation and restraints being lifted from the protein. In general, it is important for the RMSD versus time curve to equilibrate, or reach a steady state. If an RMSD curve continues to increase, it shows that the system is either still equilibrating or the protein is switching to a new conformational state. A steady curve does not, however, suggest convergence. As noticeable from the curves in Figure 4.4, the RMSD values did not change significantly from the starting structure and each curve is beginning to equilibrate. It will be important to continue running all four simulations longer to ensure the RMSD values do not increase.

Another way to qualitatively observe the protein movements over the course of a simulation is to calculate the root mean squared fluctuation (RMSF). Here, RMSF describes the fluctuations of each Cα atom of the amino acid residues in the protein averaged over the simulation time (Figure 4.5). It is clear from all four plots that the regions of the protein with the least amount of movement are the transmembrane regions, those denoted by dark grey boxes. On average, intracellular loop 3 shows the greatest fluctuations, which is expected since it is the longest loop. The N- and C-termini have been excluded from the plot due to the large fluctuations seen in these regions of the protein. Including the termini made it difficult to see the fluctuations in the other regions of the protein.

A third qualitative assessment of the simulations is shown in Figure 4.6. The figure shows heatmaps of an all-to-all RMSD calculation performed on the transmembrane Cα atoms of the Tre1 protein in mtre1, msctt, gtre1 and gsctt model systems. An all-to-all RMSD calculation is a technique used to visually see the

number of different states the protein has visited during the course of the entire simulation. The heatmap plots show block structures on the diagonal that represent the different protein conformations visited by the protein. In heatmap plots, block structures off the diagonal represent a protein conformation being re-visited at a later time during the course of the simulations. In the heatmaps of mtre1, msctt, gtre1 and gsctt, off-diagonal blocks are not present, which suggests the simulations are not well sampled.

Studies of the NRY motif of Tre1

From the genetic studies discussed in Chapter 2, it is known that the arginine of the NRY motif in Tre1 is critical to the function of this GPCR [47]. Other than the critical nature of the arginine to Tre1 function, very little is known about the arginine. It is possible that the arginine is involved in forming a salt bridge with an aspartic acid residue in TM6. A similar salt bridge in other GPCRs is thought to be important for holding GPCRs in inactive or activated states [50-60]. The hypothesis of this study was that the arginine in Tre1 is involved in forming a salt bridge with TM6 and that when Tre1 lacks this arginine ($Tre1^{sctt}$) the Tre1 GPCR has impaired function. It is possible that an alternative arginine could be used in msctt and gsctt to form a different salt bridge. This alternative salt bridge could explain why the $tre1^{sctt}$ allele does not appear to be a complete loss-of-function allele of the *tre1* gene [47]. The alternative arginine is located one residue from where the original arginine is located in $Tre1^{+}$ (Figure 4.1). Therefore, the arginine could be close enough to form an alternative salt bridge in the $Tre1^{sctt}$ protein. To test this hypothesis, the formation of a salt bridge was evaluated in all four model systems. A salt bridge is defined as a noncovalent interaction between the carboxylate group of glutamic acid or aspartic acid and the guanidium group of arginine or the ammonium of lysine. As the aspartic acid residue in TM6 has two oxygens that could be involved in a salt bridge and the arginine in TM3 has two nitrogens that could be involved in a salt bridge, the distance between both the oxygens of the aspartic acid and both the nitrogens of the arginine were calculated and plotted over the simulation time (Figure 4.7). Interatomic distances of 3.2 Å or less were considered favorable for salt bridge formation and such distances were seen in mtre1, msctt and gtre1. The atoms studied in gsctt were never close enough to form a salt bridge. Visually looking at the gsctt system, the nitrogens of ARG135 are not situated towards the oxygens of the aspartic acid residue in TM6 as they are in the mtre1, msctt and gtre1 systems. N-O

distances in mtre1 were not consistently within 3.2 Å of each other until the last ~25 ns of the simulation. gtre1, on the other hand, shows an interatomic N-O distance of 3.2 Å or less for 3 of the 4 possible N-O pairs throughout most of the simulation. It is clear that the structures of mtre1 and gtre1 are quite different in this region of Tre1 at this point in the simulation. Whether this is because mtre1 and gtre1 represent different conformations of Tre1 or because of inherent differences from the initial protein structure predictions is not yet known. msctt does not appear to be able to form a stable salt bridge using the alternative arginine, ARG135. In msctt, the correct interatomic N-O distances are transient and random throughout the simulation.

Another way to study the role of the arginine of the NRY motif in salt bridge formation is to calculate the solvent accessible surface area (SASA) around the residue. The hypothesis is that the arginine of the NRY motif is involved in a stable salt bridge in $Tre1^{+}$ and will have less SASA than the alternative arginine in $Tre1^{sctt}$ that is involved in a transient salt bridge or no salt bridge. The model is that the arginine of the NRY motif of $Tre1^{+}$ will have less SASA since the arginine and the aspartic acid residues that interact to form a salt bridge need to be close to each other. This means that TM3 and TM6 will also need to be relatively close to one another, preventing solvent from interacting with the arginine residue, which is located near the end of TM3. The results from calculating the SASA of the arginine of the NRY motif agree with this hypothesis to some extent (Figure 4.8). As discussed previously, the ability to form a stable salt bridge is present in gtre1 and it appears that a stable salt bridge is potentially forming in mtre1. On the other hand, no salt bridge is able to form in gsctt and a very transient salt bridge, if any, can form in msctt. Just like there are differences in salt bridge formation and stability, there are differences in SASA for the arginine residues in the different model systems. ARG134 is the arginine believed to be involved in salt bridge in mtre1 and gtre1. As shown, the SASA for this residue is much lower than the control arginines (Figure 4.8). The control arginines, ARG233 (ARG225 in $Tre1^{sctt}$) and ARG370 (ARG362 in $Tre1^{sctt}$) are located in intracellular loop 3 and the C-terminal tail, respectively. The control arginines represent arginines that are not in close contact with another helix in the protein. At times the control arginines are close to other loops of the protein in the cytoplasm, and that is represented by a decrease in SASA. The SASA calculations were performed on eight data points for each different arginine. Eventually, the SASA will be calculated over the course of the entire trajectory, but at this time, the data is enough to see the general trend. In contrast to the results seen with mtre1 and gtre1, msctt and

gsctt arginines have a much more difficult SASA plot to interpret. In general, the curves for the alternative arginine, ARG135, are higher than the SASA curves of the corresponding wild type structures. This increase in SASA could inhibit the alternative arginine from forming a strong interaction with the aspartic acid residue in TM6, and thus create a weaker or transient salt bridge.

Conclusions

The Tre1 GPCR is an important component of primordial germ cell migration in Drosophila [45-47]. In a partial loss-of-function allele of the *tre1* gene, *tre1*^{sctt}, proper primordial germ cell migration is disrupted. The Tre1^{sctt} protein is missing eight amino acids, RYILIACH, from the junction of the third transmembrane domain and second intracellular loop [47]. Determining how the loss of the amino acids RYILIACH affects Tre1 structure and function was the aim of this study.

Protein structure predictions were generated for Tre1^{+} and Tre1^{sctt} using GPCR-ModSim [39] and GPCR-ITASSER [40]. The four resulting structures were inserted into a POPE lipid bilayer and subjected to 200 ns of MD simulations. Although tests calculating the decorrelation times and convergence of the simulations have shown that the data generated by the 200 ns of MD is under-sampled and not converged, some insights into the structures of Tre1^{+} and Tre1^{sctt} were gained.

First, as shown by the RMSD versus time plots and the RMSF plots (Figures 4.4 and 4.5), Tre1^{+} and Tre1^{sctt} behave similarly. The RMSD values for both Tre1^{+} and Tre1^{sctt} protein structure predictions begin to stabilize near the end of the 200 ns MD run. Also, as shown by Figure 4.5, the general fluctuations of specific regions of the proteins are similar between Tre1^{+} and Tre1^{sctt}. At this qualitative level, there are greater differences between the protein structure predictions generated by GPCR-ModSim and GPCR-ITASSER than between the wild type and mutant forms of the proteins.

The only primary sequence difference between Tre1^{+} and Tre1^{sctt} is that Tre1^{sctt} is missing the eight amino acids RYILIACH. Of the eight amino acids, the arginine is the most conserved residue, being present in 96% of Rhodopsin family GPCRs [49]. The arginine and the tyrosine are also part of a highly conserved D/ERY motif in Rhodopsin family GPCRs [3,48]. The D/ERY motif is thought to have roles as a microswitch, being involved in holding the GPCR in an active or inactive state by forming a salt bridge with an aspartic acid

residue in TM6 [7,50-60]. Interestingly, the arginine is also the most critical residue of the eight amino acids for proper primordial germ cell migration in Drosophila. When the arginine is substituted for an alanine, a severe loss-of-function germ cell phenotype is observed [47]. Based upon this knowledge, it was hypothesized that a salt bridge involving the conserved D/ERY motif (NRY in Tre1) is present in $Tre1^{+}$ and allows for efficient signal transduction. This salt bridge may be missing in $Tre1^{sctt}$ and the lack of a salt bridge in $Tre1^{sctt}$ could impair proper signal transduction through Tre1. It is also possible that an alternative salt bridge using a nearby arginine could be formed in $Tre1^{sctt}$. If an alternative salt bridge forms, it could be involved in restoring some function of the GPCR. An alternative hypothesis is that the lack of the eight amino acids could affect Tre1 structure or stability so drastically that the GPCR is not folded properly and is not inserted in a membrane. This hypothesis seems less likely to be true as studies using the R33-encoded GPCR and the murine cytomegalovirus chemokine receptor homolog M33 have found that substituting the arginine of the NRY motif for an alanine residue abolishes signaling from the receptor but does not change the protein expression level [82,83]. While substituting an alanine for an arginine is not as severe as complete deletion of eight amino acids, a similar germ cell phenotype was seen in Tre1 when the eight amino acids were missing and when the arginine was substituted for an alanine (*tre1*^{sctt} reconstruction and AYILIACH in Figure 2.3) [47].

The ability of the four model systems to form a salt bridge (Figure 4.7) and the SASA around the arginine of the NRY motif (Figure 4.8) was examined. Wild type systems, mtre1 and gtre1, confirm that it is possible for a salt bridge to form between the arginine of the NRY motif (ARG134) and an aspartic acid (ASP266) in TM6. The salt bridge analysis using the mutant systems, msctt and gsctt, did not give as clear of answers. While it is possible for a salt bridge to form between the alternative arginine (ARG135) and the aspartic acid of TM6 (ASP258) in the msctt system, the salt bridge would not be very stable. Distances favorable for salt bridge formation were not consistently present during the simulation (Figure 4.7). It is clear from Figure 4.7 that no salt bridge can form in gsctt. SASA of the arginine of the NRY motif (ARG134 or ARG135) was calculated to assess the possibility of salt bridge formation (Figure 4.8). If an arginine was involved in a stable salt bridge, it is possible that the arginine would have a lower SASA than an arginine not involved in a salt bridge. The data collected from this experiment was not sufficient to arrive at any definitive conclusions. At this time it appears as if the SASA of the selected arginines (ARG134 and ARG135) are lower

in the wild type systems when compared to the mutant systems. Taken together, the salt bridge analysis and the SASA data show that it is possible for a stable salt bridge to form in the wild type forms of the protein and not likely for a stable salt bridge to form in the mutant forms of the protein. The inability to form a stable salt bridge could make the $Tre1^{sctt}$ protein structure unstable and unable to properly transduce signals.

The work presented here is important for the Drosophila primordial germ cell community as well as the GPCR community. Studying the structure of $Tre1^{+}$ and $Tre1^{sctt}$ as it relates to function is critical to a more complete understanding of the role Tre1 plays in primordial germ cell migration. Additionally, studying Tre1 and specifically the arginine of the D/ERY motif is of interest to the GPCR community since the arginine residue is highly conserved among Rhodopsin family GPCRs [49]. This study is unique in that there was an experimental observation that was the basis for the structural analysis of Tre1, and determining the structure of Tre1 when it is missing the critical arginine ($Tre1^{sctt}$) will provide the GPCR community with more information about the arginine of the D/ERY motif in a different context than has been previously studied.

References

1. Fredriksson R, Lagerström MC, Lundin L-G, Schiöth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Molecular pharmacology 63: 1256-1272.
2. Gloriam DE, Fredriksson R, Schiöth HB (2007) The G protein-coupled receptor subset of the rat genome. BMC Genomics 8: 338.
3. Lagerström MC, Schiöth HB (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. Nature Reviews Drug Discovery 7: 339-357.
4. Cherezov V, Abola E, Stevens RC (2010) Recent progress in the structure determination of GPCRs, a membrane protein family with high potential as pharmaceutical targets. Methods in molecular biology 654: 141-168.
5. Xu F, Liu W, Hanson MA, Stevens RC, Cherezov V (2011) Development of an Automated High Throughput LCP-FRAP Assay to Guide Membrane Protein Crystallization in Lipid Mesophases. Crystal Growth & Design 11: 1193-1201.
6. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, et al. (2007) High-Resolution Crystal Structure of an Engineered Human 2-Adrenergic G Protein Coupled Receptor. Science 318: 1258-1265.
7. Chien EYT, Liu W, Zhao Q, Katritch V, Han GW, et al. (2010) Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. Science 330: 1091-1095.
8. Jaakola V-P, Griffith MT, Hanson MA, Cherezov V, Chien EYT, et al. (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 322: 1211-1217.
9. Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, et al. (2010) Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. Science 330: 1066-1071.
10. Palczewski K, Kumasaka T, Hori T, Behnke C, Motoshima H, et al. (2000) Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. Science 289: 739-745.
11. Rasmussen S, Choi H, Rosenbaum D, Kobilka T, Thian F, et al. (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450: 383-388.

12. Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, et al. (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. Nature 454: 486-491.
13. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, et al. (2011) Structure of the human histamine H1 receptor complex with doxepin. Nature 475: 65-70.
14. Haga K, Kruse AC, Asada H, Yurugi-Kobayashi T, Shiroishi M, et al. (2012) Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. Nature 482: 23-28.
15. Kruse AC, Hu J, Pan AC, Arlow DH, Rosenbaum DM, et al. (2012) Structure and dynamics of the M3 muscarinic acetylcholine receptor. Nature 482: 552-556.
16. Hanson MA, Roth CB, Jo E, Griffith MT, Scott FL, et al. (2012) Crystal Structure of a Lipid G Protein-Coupled Receptor. Science 335: 851-855.
17. Wu H, Wacker D, Mileni M, Katritch V, Han GW, et al. (2012) Structure of the human κ-opioid receptor in complex with JDTic. Nature: 1-8.
18. Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, et al. (2012) Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. Nature: 1-7.
19. Granier S, Manglik A, Kruse AC, Kobilka TS, Thian FS, et al. (2012) Structure of the δ-opioid receptor bound to naltrindole. Nature 485: 400-404.
20. Thompson AA, Liu W, Chun E, Katritch V, Wu H, et al. (2012) Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic. Nature 485: 395-399.
21. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, et al. (2008) Crystal structure of opsin in its G-protein-interacting conformation. Nature 455: 497-502.
22. Choe H-W, Kim YJ, Park JH, Morizumi T, Pai EF, et al. (2011) Crystal structure of metarhodopsin II. Nature 471: 651-655.
23. Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, et al. (2011) Crystal structure of the β2 adrenergic receptor-Gs protein complex. Nature 477: 549-555.
24. Park JH, Scheerer P, Hofmann KP, Choe H-W, Ernst OP (2008) Crystal structure of the ligand-free G-protein-coupled receptor opsin. Nature 454: 183-187.
25. Standfuss J, Edwards PC, D'Antona A, Fransen M, Xie G, et al. (2011) The structural basis of agonist-induced activation in constitutively active rhodopsin. Nature 471: 656-660.
26. Xu F, Wu H, Katritch V, Han GW, Jacobson KA, et al. (2011) Structure of an agonist-bound human A2A adenosine receptor. Science 332: 322-327.
27. Baker D, Sali A (2001) Protein structure prediction and structural genomics. Science 294: 93-96.
28. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, et al. (2009) Protein structure homology modeling using SWISS-MODEL workspace. Nature Protocols 4: 1-14.
29. Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature Protocols 5: 725-738.
30. Zhang Y, Devries ME, Skolnick J (2006) Structure modeling of all identified G protein-coupled receptors in the human genome. PLoS computational biology 2: e13.
31. Worth C, Kleinau G, Krause G (2009) Comparitive sequence and structural analyses of G-protein-coupled-receptor crystal structures and implications for molecular models. PLoS ONE 4: 1-14.
32. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25: 3389-3402.
33. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. Journal of Molecular Biology 292: 195-202.
34. Mobarec JC, Sanchez R, Filizola M (2009) Modern Homology Modeling of G-Protein Coupled Receptors: Which Structural Template to Use? Journal of Medicinal Chemistry 52: 5207-5216.
35. Sokkar P, Mohandass S, Ramachandran M (2010) Multiple templates-based homology modeling enhances structure quality of AT1 receptor: validation by molecular dynamics and antagonist docking. Journal of Molecular Modeling 17: 1565-1577.
36. Zhang Y (2008) Progress and challenges in protein structure prediction. Current Opinion in Structural Biology 18: 342-348.
37. Zhang Y (2009) Protein structure prediction: when is it useful? Current Opinion in Structural Biology 19: 145-155.
38. Worth CL, Kreuchwig A, Kleinau G, Krause G (2011) GPCR-SSFE: A comprehensive database of G-protein-coupled receptor template predictions and homology models. BMC Bioinformatics 12: 185.

39. Rodríguez D, Bello X, Gutiérrez-de-Terán H (2012) Molecular Modelling of G Protein-Coupled Receptors Through the Web. Molecular Informatics.
40. Zhang J, Zhang Y (2010) GPCRRD: G protein-coupled receptor spatial restraint database for 3D structure modeling and function annotation. Bioinformatics 26: 3004-3005.
41. Zhang J, Liang Y, Zhang Y (2011) Atomic-Level Protein Structure Refinement Using Fragment-Guided Molecular Dynamics Conformation Sampling. Structure 19: 1784-1795.
42. Zhang J, Zhang Y (2010) A Novel Side-Chain Orientation Dependent Potential Derived from Random-Walk Reference State for Protein Fold Selection and Structure Prediction. PLoS ONE 5: e15386.
43. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, et al. (2000) Comparative protein structure modeling of genes and genomes. Annual Review of Biophysics and Biomolecular Structure 29: 291-325.
44. Eswar N, Eramian D, Webb B, Shen M-Y, Sali A (2008) Protein structure modeling with MODELLER. Methods in molecular biology 426: 145-159.
45. Kunwar P, Starz-Gaiano M, Bainton R, Heberlein U, Lehmann R (2003) Tre1, a GPCR, directs transepithelial migration of Drosophila germ cells. Plos Biol 1: 372-384.
46. Kunwar PS, Sano H, Renault AD, Barbosa V, Fuse N, et al. (2008) Tre1 GPCR initiates germ cell transepithelial migration by regulating Drosophila melanogaster E-cadherin. J Cell Biol 183: 1-12.
47. Kamps AR, Pruitt MM, Herriges JC, Coffman CR (2010) An Evolutionarily Conserved Arginine is Essential for Tre1 G Protein-Coupled Receptor Function During Germ Cell Migration in Drosophila melanogaster. PLoS ONE 5: e11839.
48. Nygaard R, Frimurer TM, Holst B, Rosenkilde MM, Schwartz TW (2009) Ligand binding and micro-switches in 7TM receptor structures. Trends in Pharmacological Sciences 30: 249-259.
49. Mirzadegan T, Benkö G, Filipek S, Palczewski K (2003) Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. Biochemistry 42: 2759-2767.
50. Aizaki Y, Maruyama K, Nakano-Tetsuka M, Saito Y (2009) Distinct roles of the DRY motif in rat melanin-concentrating hormone receptor 1 in signaling control. Peptides 30: 974-981.
51. Ballesteros J, Kitanovic S, Guarnieri F, Davies P, Fromme BJ, et al. (1998) Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. J Biol Chem 273: 10445-10453.
52. Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, et al. (2001) Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. The Journal of Biological Chemistry 276: 29171-29177.
53. Shapiro DA, Kristiansen K, Weiner DM, Kroeze WK, Roth BL (2002) Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. The Journal of Biological Chemistry 277: 11441-11449.
54. Vogel R, Mahalingam M, Ludeke S, Huber T, Siebert F, et al. (2008) Functional Role of the "Ionic Lock" - An Interhelical Hydrogen-Bond Network in Family A Heptahelical Receptors. Journal of Molecular Biology 360: 648-655.
55. Zhu SZ, Wang SZ, Hu J, el-Fakahany EE (1994) An arginine residue conserved in most G protein-coupled receptors is essential for the function of the m1 muscarinic receptor. Molecular pharmacology 45: 517-523.
56. Scheer A, Costa T, Fanelli F, DeBenedetti P, Mhaouty-Kodja S, et al. (2000) Mutational Analysis of the Highly Conserved Arginine within the Glu/Asp-Arg-Tyr Motif of the alpha1b-Adrenergic Receptor: Effects on Receptor Isomerization and Activation. Molecular pharmacology 57: 219-231.
57. Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. EMBO J 15: 3566-3578.
58. Greasley PJ, Fanelli F, Rossier O, Abuin L, Cotecchia S (2002) Mutagenesis and modelling of the alpha(1b)-adrenergic receptor highlight the role of the helix 3/helix 6 interface in receptor activation. Molecular pharmacology 61: 1025-1032.
59. Capra V, Veltri A, Foglia C, Crimaldi L, Habib A, et al. (2004) Mutational analysis of the highly conserved ERY motif of the thromboxane A2 receptor: alternative role in G protein-coupled receptor signaling. Molecular pharmacology 66: 880-889.

60. Greasley PJ, Fanelli F, Scheer A, Abuin L, Nenniger-Tosato M, et al. (2001) Mutational and Computational Analysis of the alpha1b-Adrenergic Receptor: Involvement of Basic and Hydrophobic Residues in Receptor Activation and G Protein Coupling. The Journal of Biological Chemistry 276: 46485-46494.
61. Fanelli F, De Benedetti PG (2011) Computational Modeling Approaches to Structure−Function Analysis of G Protein-Coupled Receptors. Chemical Reviews 111: PR438-PR535.
62. Phillips J, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable Molecular Dynamics with NAMD. Journal of Computational Chemistry 26: 1781-1802.
63. Jo S, Kim T, Im W (2007) Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. PLoS ONE 2: e880.
64. Jo S, Lim JB, Klauda JB, Im W (2009) CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. Biophysj 97: 50-58.
65. Jones H, Harwood J, Bowen I, Griffiths G (1992) Lipid Composition of Subcellular Membranes from Larvae and Prepupae of Drosophila melanogaster. LIPIDS 27: 984-987.
66. MacKerell A, Bashford D, Bellott M, Dunbrack R, Evanseck J, et al. (1998) All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. Journal of Physical Chemistry B 102: 3586-3616.
67. MacKerell Jr AD (2004) Empirical Force Fields for Biological Macromolecules: Overview and Issues. Journal of Computational Chemistry 25: 1584-1604.
68. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, et al. (2010) Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. Journal of Physical Chemistry B 114: 7830-7843.
69. Jorgensen W, Chandrasekhar J, Madura J, Impey R, Klein M (1983) Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics 79: 926-935.
70. Feller S, Zhang Y, Pastor R, Brooks B (1995) Constant pressure molecular dynamics simulation: The Langevin piston method. Journal of Chemical Physics 103: 4613-4621.
71. Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. The Journal of Chemical Physics 81: 511.
72. Darden T, York D, Pedersen L (1992) Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. The Journal of Chemical Physics 98: 10089-10092.
73. Humphry W, Dalke A, Schulten K (1996) VMD: Visual Molecular Dynamics. Journal of Molecular Graphics 14: 33-38.
74. Romo T, Grossfield A (2009) LOOS: An Extensible Platform for the Structural Analysis of Simulations. Conference of the IEEE EMBS: 2332-2335.
75. Lyman E, Zuckerman DM (2006) Ensemble-Based Convergence Analysis of Biomolecular Trajectories. Biophysical Journal 91: 164-172.
76. Lyman E, Zuckerman DM (2007) On the structural convergence of biomolecular simulations by determination of the effective sample size. The Journal of Physical Chemistry B 111: 12876-12882.
77. Zhang X, Bhatt D, Zuckerman DM (2010) Automated Sampling Assessment for Molecular Simulations Using the Effective Sample Size. Journal of Chemical Theory and Computation 6: 3048-3057.
78. Romo TD, Grossfield A (2011) Block Covariance Overlap Method and Convergence in Molecular Dynamics Simulation. Journal of Chemical Theory and Computation 7: 2464-2472.
79. Mori T, Ogushi F, Sugita Y (2011) Analysis of lipid surface area in protein-membrane systems combining voronoi tessellation and monte carlo integration methods. Journal of Computational Chemistry.
80. Rappolt M, Hickel A, Bringezu F, Lohner K (2003) Mechanism of the Lamellar:Inverse Hexagonal Phase Transition Examined by High Resolution X-ray Diffraction. Biophysical Journal 84: 3111-3122.
81. Grossfield A, Zuckerman DM (2009) Quantifying uncertainty and sampling quality in biomolecular simulations. Annual reports in computational chemistry 5: 23-48.
82. Gruijthuijsen YK, Beuken E, Smit M, Leurs R, Bruggeman C, et al. (2004) Mutational analysis of the R33-encoded G protein-coupled receptor of rat cytomegalovirus: identification of amino acid residues critical for cellular localization and ligand-independent signalling. Journal of General Virology 85: 897-909.
83. Case R, Sharp E, Benned-Jensen T, Rosenkilde MM, Davis-Poynter N, et al. (2008) Functional Analysis of the Murine Cytomegalovirus Chemokine Receptor Homologue M33: Ablation of Constitutive Signaling Is Associated with an Attenuated Phenotype In Vivo. Journal of Virology 82: 1884-1898.

CHAPTER 5. GENERAL CONCLUSIONS

An evolutionarily conserved arginine in the Tre1 G protein-coupled receptor is critical for proper primordial germ cell migration

An EMS mutagenesis screen, aimed at finding genes involved in primordial germ cell migration and programmed cell death, resulted in the identification of the *scattershot* gene [1]. The primordial germ cells in *scattershot* mutant embryos did not migrate properly to the gonads, instead "scattering" across the posterior half of the embryo. In 2003, Kunwar et al. showed that *scattershot* failed to complement $tre1^{\Delta EP5}$, a mutant allele of the gene that encodes the Tre1 G protein-coupled receptor (GPCR), suggesting that *scattershot* may be an allele of *tre1*. However, Kunwar et al. reported that there were no disruptions to the *tre1* coding region in *scattershot*, making it unclear whether the failure to complement was allelic non-complementation or non-allelic non-complementation [2]. Chapter 2 describes the work performed that conclusively shows *scattershot* is an allele of *tre1*.

The introns and coding regions of *tre1* were sequenced and revealed a single significant base pair change in *scattershot* flies. An adenine was switched to a thymine at the junction of intron 4 and exon 5 that resulted in a change from an AG to a TG, altering the usual splice acceptor site. Sequencing of $tre1^{sctt}$ cDNA confirmed that a cryptic splice acceptor site 24 base pairs 3' was utilized instead in $tre1^{sctt}$ flies. The change from an AG to a TG resulted in an in-frame deletion of eight amino acids, RYILIACH [3].

A systematic approach was taken to determine if the eight missing amino acids or a subset of these amino acids was critical for proper primordial germ cell migration. Transgenic lines containing specific alanine substitutions of individual amino acids or subsets of these eight amino acids, as well as deletion of these eight amino acids, were constructed in a $tre1^{sctt}$ background. The maternal rescue experiments detailed in Chapter 2 illustrate the critical role of the arginine in proper primordial germ cell migration. When an alanine is substituted for the arginine (AYILIACH), the construct fails to rescue primordial germ cell migration, and the embryos show a completely *scattershot*-like phenotype, with 0.4 or 1.1 primordial germ cells migrating properly to the gonads (Table 2.2) [3]. The arginine is part of the highly conserved D/ERY motif that is important for proper signal transduction in many GPCRs [4-13].

After completing the maternal rescue experiments, an outstanding question was whether the transgenes could rescue the *tre1*^{sctt} mutant phenotype when expressed paternally in a *tre1*^{sctt} mutant background. When a homozygous *tre1*^{sctt} female is crossed to a wild-type male, the paternal rescue phenotype is observed in half of the embryos, specifically the *tre1*^{sctt} /+ females (Figure 2.1 and Figure 5.1) [1,3]. In the paternal rescue phenotype, wild type numbers of primordial germ cells (more than 12 primordial germ cells) are in the gonads and some primordial germ cells are left ectopic to the gonads. To test if zygotic expression of the transgenes is sufficient to rescue primordial germ cell migration, homozygous *tre1*^{sctt} females were crossed to *tre1*^{sctt} males carrying two copies of the transgene, one on each of the two second chromosomes. It was hypothesized that when fully functional transgenes are supplied paternally, the transgenes will rescue in a similar fashion as wild-type males do. 100% of the progeny are expected to rescue migration, since each embryo will carry one copy of the paternally supplied transgene. The positive control cross (T^+G^+ in a *tre1*^{sctt} background, named RYILIACH) and the negative control crosses (no transgene and *tre1*^{sctt} reconstruction) behaved as expected (Figure 5.1 and Table 5.1). When RYILIACH is supplied paternally, on average 16.2 primordial germ cells migrate properly to the gonads. Comparatively, when the no transgene control and *tre1*^{sctt} reconstruction (two different insertion sites) are supplied paternally, an average of 1.2, 1.0 and 1.1 primordial germ cells migrate to the gonads, respectively. Similar to the maternal rescue experiments, the paternal rescue experiments show the tyrosine is not critical for Tre1 function. On average, 12.5 and 12.2 primordial germ cells migrate properly to the gonads when the tyrosine is replaced with an alanine (RAILIACH, Table 5.1). The results also agree with the maternal rescue experiments in that when the arginine is switched to an alanine (AYILIACH), primordial germ cell migration is not rescued (Figure 5.1 and Table 5.1). When the transgene AYILIACH is supplied paternally, 0.7 and 1.0 primordial germ cells migrate properly to the gonads on average, with all of the remaining primordial germ cells left in ectopic positions. The failure of the AYILIACH construct to rescue when supplied paternally shows, as do the results from the maternal rescue experiments, that the arginine is critical for proper primordial germ cell migration and function of Tre1.

The remaining transgenes, RYAAAAAA and RYAAAACH, provided interesting results. The RYAAAAAA and RYAAAACH constructs rescued primordial germ cell migration to wild type levels when supplied maternally [3] but only partially rescued migration when supplied paternally. With both of these

transgenes, two phenotypes are observed: the first phenotype shows the *scattershot* phenotype (no rescue of migration), the second phenotype shows partial rescue of migration (Figure 5.1). The second phenotype is described as a partial rescue of migration since the number of germ cells in the gonads is reduced compared to wild type paternal rescue. Specifically, an embryo is classified as having a partial rescue of migration phenotype when the total number of primordial germ cells in the gonads is greater than four. For the RYAAAACH transgene, 40% of the embryos display the partial rescue of migration phenotype with an average of 8.2 primordial germ cells in the gonads (Table 5.1). Two independent insertion sites were tested for the RYAAAAAA transgene. In the maternal rescue experiment, the two RYAAAAAA independent insertion lines behaved similarly (Figure 2.3) [3], yet when tested paternally they did not. RYAAAAAA (1) has 76% of the embryos displaying the partial rescue of migration phenotype, while RYAAAAAA (2) has 40%. Additionally, the number of primordial germ cells that migrate properly to the gonads differed between the two transgene insertions. RYAAAAAA (1) has an average of 7.1 primordial germ cells in the gonads while the RYAAAAAA (2) line rescues migration more similar to wild type with an average of 13.2 primordial germ cells in the gonads (Table 5.1).

It is interesting that when RYAAAACH and RYAAAAAA are supplied paternally, only a partial rescue of primordial germ cell migration is observed. Even more intriguing is that in the embryos where primordial germ cell migration is rescued, the number of primordial germ cells in the gonads is much lower than wild type. The results suggest that perhaps the timing of expression of the transgenes is delayed. Zygotic transcription begins at stages 8-9 [14], so it is possible that RYAAAACH or RYAAAAAA are not transcribed early enough to rescue the primordial germ cell migration defect. Another possibility is that there is a dependence on dosage of gene function and one copy of the RYAAAACH or RYAAAAAA transgenes supplied paternally is not sufficient to rescue primordial germ cell migration.

To test if the partial rescue of migration is caused by a dosage effect, recombinant lines carrying the two RYAAAAAA transgenes on a single chromosome were made. Three independent recombinant lines were constructed and were designated RYAAAAAA recomb 1, 2 and 3. Primordial germ cell migration is rescued to nearly wild type levels in close to 100% of the embryos when the recombinant lines are used to supply two copies of the transgene paternally (Figure 5.1). On average, 10.9, 11.8 and 11.1 primordial germ cells migrate

properly to the gonads in RYAAAAAA recomb 1, 2 and 3, respectively (Table 5.1). These results suggest that the failure of RYAAAAAA to rescue primordial germ cell migration when supplied paternally is due to a dosage effect.

The results from the RYAAAAAA recombinant lines reveal a role of the amino acids ILIACH in Tre1 function. When RYAAAAAA is supplied maternally, primordial germ cell migration is rescued, but 8.5 and 8.9 primordial germ cells, on average, remain in ectopic locations [3]. In contrast, when RYILIACH is supplied maternally, an average of 1.2 primordial germ cells persist in ectopic locations. More ectopic primordial germ cells are present when RYAAAAAA or the RYAAAAAA recombinant lines with two copies of the transgene are supplied paternally than when RYILIACH is supplied paternally in a single copy dose (Table 5.1). Additional constructs, specifically targeting the amino acids ILIACH could be made to further investigate the role of ILIACH in Tre1 function. Specifically, the RAAAIAAA construct would be an interesting choice. The isoleucine not substituted to alanine in RAAAIAAA is part of a microdomain around the DRY motif in GPCRs, (I/L)xxDRYxx(I/L) [4]. Results with the gonadotropin-releasing hormone receptor have shown that substituting the isoleucine located two amino acids carbonyl to the DRY motif to alanine allowed the arginine of the DRY motif to accept a new conformational state. In this new conformation, the arginine was more exposed to the cytoplasm and was thus more heavily solvated. The extra solvation prevented the arginine from forming specific intramolecular interactions necessary for establishing an active receptor state [4]. The results seen with both the maternal and paternal rescue experiments using RYAAAACH and RYAAAAAA are consistent with this model. It is also possible that the amino acids ILIACH are required for binding of specific proteins to the DRY motif. Substituting ILI or ILIACH with much smaller alanine residues may block the optimal binding of a protein to the DRY motif and therefore inhibit the proper signaling response. Other transgenic lines could be constructed to test which of the amino acids ILIACH are critical for eliminating the ectopic primordial germ cells.

As described previously, when an embryo has double the amount of the RYAAAAAA transgene (the RYAAAAAA recombinant lines), primordial germ cell migration is rescued to nearly wild type levels. In these embryos, as well as those from the single copy paternal rescue experiments, primordial germ cell migration can be rescued by supplying a functional copy of *tre1*, but there are primordial germ cells remaining in ectopic

positions (Table 5.1). To test the hypothesis that additional doses of *tre1* will eliminate the ectopic primordial germ cells when expressed paternally, one could cross $tre1^{sctt}$ virgins to T^+G^+ males. The T^+G^+ transgene contains the sequence of *tre1* and *Gr5a*, a neighboring gene, and was used to build the transgenic lines described in Chapter 2 [3,15]. Therefore, T^+G^+ males carry the wild-type *tre1* gene on the X chromosome and the T^+G^+ transgene on the second chromosome. It would be interesting to see if more than one copy of wild-type *tre1* supplied paternally could eliminate the ectopic primordial germ cells, completely restoring proper primordial germ cell development. To further analyze the effects of transgene dosage, a cross could be performed between an RYAAAAAA recombinant female and a $tre1^{sctt}$ male. When RYAAAAAA is supplied maternally, primordial germ cell migration is rescued, but many primordial germ cells remain in ectopic positions (Figure 2.3). It is possible that an extra copy of RYAAAAAA on each second chromosome supplied maternally could eliminate the ectopic primordial germ cells.

From studies using the complete loss-of-function allele of *tre1*, $tre1^{\Delta EP5}$, it was shown that Tre1 plays roles in the polarization, dispersal and transepithelial migration of the primordial germ cells [16]. If multiple copies of *tre1* supplied paternally do not eliminate the ectopic primordial germ cells, it suggests that Tre1 could have additional roles in mediating primordial germ cell development. Studies with $tre1^{\Delta EP5}$ showed that the primordial germ cells in these mutants do not exit the posterior midgut. Double staining $tre1^{sctt}$ embryos with a primordial germ cell marker and a midgut marker did not allow the determination of whether or not all of the primordial germ cells exited the midgut in $tre1^{sctt}$ embryos due to morphological rearrangements of the posterior midgut epithelium (Figure 3.2). Since that work, the lab has obtained lines in which three types of red fluorophores are expressed in the germ cells, mCherry-Vasa [17], nanos-mRFPruby-moesin [18] and nanos-mKate2-moesin [19]. Live imaging with these lines in various *tre1* backgrounds could be used to further study this time point in development and determine whether the primordial germ cells exit the posterior midgut and if so, under what conditions. Once this is established, it could be interesting to investigate whether the primordial germ cells exit the posterior midgut when the transgenes are expressed maternally and paternally.

Signaling networks governing Drosophila primordial germ cell migration

The primordial germ cells of Drosophila have been used as a model for cell migration for many years.

While some proteins mediating Drosophila primordial germ cell migration have been identified, an understanding of the signaling networks involved has remained incomplete. In Chapter 3, a genetic loss-of-function screen and a yeast two-hybrid screen were performed to identify additional proteins involved in Drosophila primordial germ cell migration. From both screens, two lipid phosphate phosphatases, Wunen (Wun) and Wunen2 (Wun2), were identified as potential components of the Tre1 signaling network.

Tre1, Wun and Wun2 are all critical to Drosophila primordial germ cell development. Tre1 has roles in the polarization, individualization and transepithelial migration of the primordial germ cells [2,16]. Wun and Wun2 are involved in the transepithelial migration of the primordial germ cells, as well as their bilateral segregation and survival [20-23]. The germ cell phenotypes of $tre1^{sctt}$ and zygotic loss-of-function *wun* and *wun2* mutants (hereafter referred to as *wun wun2*) were examined at stages 15-16 of development. In $tre1^{sctt}$ and *wun wun2* mutant embryos, proper primordial germ cell migration is disrupted, with few primordial germ cells migrating properly to the gonads (Figure 3.1 and Table 3.2). In addition to the examination of stages 15-16 germ cell phenotypes, a yeast two-hybrid screen was performed and showed that Tre1 interacts with both Wun and Wun2 (Figure 3.3). From the results of both the loss-of-function and yeast two-hybrid screens the following hypothesis was suggested: Tre1, Wun and Wun2 are involved in a common signaling network.

If *tre1*, *wun* and *wun2* are all involved in a common signaling pathway, the germ cell phenotypes at stages in development earlier than stages 15-16 should also be the same. The earlier point in development chosen to study was the first active phase of primordial germ cell migration, the exit from the posterior midgut (Figure 3.2, Table 3.3). Upon inspection of the germ cell phenotypes at stages 10-11, it appears that *wun wun2* mutant embryos do not have similar germ cell phenotypes to $tre1^{sctt}$. In control embryos, the primordial germ cells begin their exit from the posterior midgut at stage 10 and have all exited the midgut by stage 11. In *wun wun2* mutants, the primordial germ cells exit the posterior midgut at the correct stage of development, similar to the primordial germ cells in control embryos. Conversely, the exit of primordial germ cells from the posterior midgut is delayed in $tre1^{sctt}$ embryos, and whether or not all of the primordial germ cells eventually exit is unknown. Additionally, further testing of the interaction between Tre1, Wun and Wun2 in yeast by reversing the baits and the preys does not show a significant interaction (data not shown). Six additional bait constructs (three Wun and three Wun2) have been designed and are currently being tested for an interaction with Tre1.

If the additional Wun and Wun2 bait constructs show an interaction with Tre1, it confirms the initial interaction when Tre1 was the bait (Figure 3.3) and suggests that Tre1, Wun and Wun2 signaling pathways are closely coupled. At this point, an interaction between Tre1, Wun and Wun2 has been shown in yeast, however it would be beneficial to show a genetic interaction. A triple mutant using $tre1^{sctt}$ and *wun wun2* alleles could be made and the germ cell phenotypes at stages 10-11 examined. It is possible that a mutant *tre1*, $tre1^{sctt}$, in a *wun wun2* zygotic loss-of-function background could prevent the primordial germ cells from exiting the midgut at the correct stage in development. Furthermore, the ability of the primordial germ cells in $tre1^{sctt}$ to respond to *wun* and *wun2* signals could be examined. If it is determined that the primordial germ cells do exit the midgut in $tre1^{sctt}$, *wun* and *wun2* could be overexpressed in the CNS in a $tre1^{sctt}$ mutant background. Overexpression of *wun2* in the CNS in a *wun wun2* mutant embryo is able to rescue the bilateral segregation of the primordial germ cells [22]. Therefore, it is possible that the overexpression of *wun* or *wun2* in the CNS could also rescue the bilateral segregation of the primordial germ cells in a $tre1^{sctt}$ mutant embryo.

If the additional Wun and Wun2 bait constructs do not show an interaction between Tre1, Wun and Wun2, it suggests that the proposed hypothesis was wrong and that Tre1, Wun and Wun2 do not act together in a common signaling network. If this is the case, further study of the involvement of *wun* and *wun2* in germ cell death is suggested. It has been shown that on average 50% of the primordial germ cells die during stages 10-12 in Drosophila [24]. In agreement with that study, approximately 50% of the primordial germ cells die in faf^{X} (control) and $tre1^{sctt}$ embryos (Table 3.2 and Table 3.4). As discussed in Yamada et al., part of the primordial germ cell death is due to signaling pathways involving the monocarboxylate transporter, *outsiders* (*out*), and *p53* [24]. In *out* and *p53* mutants, only 25% of the primordial germ cells die and bilateral segregation is mostly normal [24]. Interestingly, essentially no germ cell death is seen in *wun wun2* zygotic loss-of-function mutants (Table 3.2 and 3.4). Based on this apparent connection, a new hypothesis for primordial germ cell death in Drosophila is proposed: 25% of the primordial germ cell death is regulated by *wun* and *wun2* expression at the midline and the other 25% is regulated by alternative mechanisms involving both *out* and *p53*. The new hypothesis suggests that *out* and *p53* pathways are either not active or not fully active in *wun wun2* zygotic loss-of-function mutants. To test this new hypothesis, *out* and/or *p53* could be overexpressed in the primordial germ cells in a *wun wun2* mutant embryo and the ability of the overexpression to rescue part of the germ cell death

phenotype examined. Additionally, according to the new hypothesis, *wun* and *wun2* pathways are functional in *out* and *p53* mutants. To confirm the functionality of the *wun* and *wun2* pathways, *wun* and *wun2* could be overexpressed in the CNS in an *out* and/or *p53* mutant. If the new hypothesis is correct, no change to the *out* or *p53* germ cell phenotype when *wun* and *wun2* are overexpressed is expected. These experiments are needed to obtain a basic understanding of the relationship between *wun*, *wun2*, *out* and *p53* in mediating primordial germ cell death.

Structural analysis of the Tre1 G protein-coupled receptor

Three-dimensional structural predictions were generated for wild type Tre1 ($Tre1^{+}$) and the mutant form of Tre1 missing the eight amino acids, RYILIACH ($Tre1^{sctt}$). The structural predictions were inserted into a POPE lipid bilayer for molecular dynamics (MD) simulations, which allowed for $Tre1^{+}$ and $Tre1^{sctt}$ to be studied in an environment that mimics the natural environment of this GPCR. The computational study of Tre1 through the use of MD simulations was performed to aid in studying the relationship between Tre1 structure and function and to gain a more comprehensive understanding of how Tre1 mediates primordial germ cell migration. The hypothesis of the study was that the eight amino acids RYILIACH are required to maintain Tre1 in a fully functional conformation. The arginine and tyrosine (RY) of these eight amino acids are part of a conserved D/ERY motif in Rhodopsin family GPCRs [4-13]. The arginine is of particular interest since it is the most conserved residue of the D/ERY motif and the arginine of the D/ERY motif can form a salt bridge with an aspartic acid or glutamic acid in transmembrane helix 6 (TM6) [25]. It is possible that a salt bridge critical in the stabilization of Tre1 forms between the arginine of the D/ERY motif (NRY motif in Tre1) and an aspartic acid residue in TM6. If this were the case, $Tre1^{sctt}$ would be inherently unstable since $Tre1^{sctt}$ is missing this specific arginine.

MD simulations were run for approximately 200 ns for each model system studied. While it is clear that the simulations have not run a sufficient amount of time to perform statistical analyses on the data (Table 4.2 and Table 4.3), a general trend in salt bridge formation between the arginine of the D/ERY motif and the aspartic acid of TM6 can be established. The data suggests that a salt bridge between the arginine of the D/ERY motif and the aspartic acid of TM6 can form in $Tre1^{+}$ (Figure 4.7). In contrast, it appears that the

alternative arginine (Figure 4.1) and the aspartic acid of TM6 cannot form a stable salt bridge in Tre1^{sctt}.

While the simulations have run for a significant amount of time, longer simulation times are necessary for a more thorough analysis of the data and testing of the hypothesis. Each of the simulations will continue running until it is determined by the decorrelation times and convergence analyses that the simulations can stop. Even though it is not possible to predict the length the simulations will need to be, continual monitoring of the decorrelation times and convergence of the simulations will aid in determining when meaningful statistical analyses can be performed. Statistical analyses on the data will allow for the role the potential salt bridge between the arginine of the D/ERY motif and the aspartic acid of TM6 plays in Tre1 structure and function to be stated with more certainty.

Additionally, once a more complete understanding of Tre1^{+} structure is obtained, MD simulations on other constructs is possible. For example, the role of the positive charge on arginine could be examined by performing MD simulations on either KYILIACH or AYILIACH. If the positive charge of the arginine of the D/ERY motif is the most critical feature of the motif, there should be minor changes to Tre1 structure in the lysine substitution model (KYILIACH). Substituting the arginine for alanine (AYILIACH) will not only test the importance of the positive charge of arginine, but the role the arginine itself plays in Tre1 structure. As described in Chapter 2, when the construct AYILIACH is expressed maternally in a *tre1^{sctt}* mutant background, the primordial germ cells fail to migrate properly to the gonads. This experiment showed that the arginine is critical to Tre1 function. It would be interesting to perform MD simulations on a Tre1 mutant containing the arginine to alanine substitution (AYILIACH) and investigate if this substitution also affects Tre1 structure. While it is possible that the arginine to alanine substitution (AYILIACH) will effect Tre1 structure, the results from the maternal rescue experiments in Chapter 2 and the paternal rescue experiments described earlier in this chapter suggest that the six amino acids ILIACH may also play important roles in Tre1 function. The construct RAAAIAAA could be tested experimentally (as suggested earlier in this chapter) or through the use of modeling to test the "arginine cage" hypothesis. The "arginine cage" hypothesis suggests that the isoleucine, located three amino acids away from the arginine of the D/ERY motif, may have roles in preventing the solvation of the arginine residue [4]. If the arginine of the D/ERY motif in Tre1^{+} were involved in forming a stable salt bridge, it would be interesting to test if this particular isoleucine is involved.

A comprehensive study of the role Tre1 plays in Drosophila primordial germ cell migration requires a detailed understanding of the nature of the GPCR. The work performed for this dissertation has provided both the primordial germ cell community and the GPCR community with a greater knowledge of the Tre1 GPCR. The determination of which amino acids are critical for proper function of the Tre1 GPCR will enable future studies on dissecting the downstream signaling networks governing Drosophila primordial germ cell migration. Additionally, the relative structure of Tre1 and Tre1 mutant proteins will aid in not only determining why specific amino acids play critical roles in Tre1 function, but could also provide a baseline for searching for potential ligands for the Tre1 GPCR.

References

1. Coffman CR, Strohm R, Oakley F, Yamada Y, Przychodzin D, et al. (2002) Identification of X-linked genes required for migration and programmed cell death of Drosophila melanogaster germ cells. Genetics 162: 273-284.
2. Kunwar P, Starz-Gaiano M, Bainton R, Heberlein U, Lehmann R (2003) Tre1, a GPCR, directs transepithelial migration of Drosophila germ cells. Plos Biol 1: 372-384.
3. Kamps AR, Pruitt MM, Herriges JC, Coffman CR (2010) An Evolutionarily Conserved Arginine is Essential for Tre1 G Protein-Coupled Receptor Function During Germ Cell Migration in Drosophila melanogaster. PLoS ONE 5: e11839.
4. Ballesteros J, Kitanovic S, Guarnieri F, Davies P, Fromme BJ, et al. (1998) Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. J Biol Chem 273: 10445-10453.
5. Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, et al. (2001) Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. The Journal of Biological Chemistry 276: 29171-29177.
6. Flanagan CA (2005) A GPCR that is not "DRY". Mol Pharmacol 68: 1-3.
7. Romo TD, Grossfield A, Pitman MC (2010) Concerted interconversion between ionic lock substates of the beta(2) adrenergic receptor revealed by microsecond timescale molecular dynamics. Biophysical Journal 98: 76-84.
8. Rosenkilde MM, Kledal TN, Schwartz TW (2005) High constitutive activity of a virus-encoded seven transmembrane receptor in the absence of the conserved DRY motif (Asp-Arg-Tyr) in transmembrane helix 3. Mol Pharmacol 68: 11-19.
9. Scheer A, Costa T, Fanelli F, DeBenedetti P, Mhaouty-Kodja S, et al. (2000) Mutational Analysis of the Highly Conserved Arginine within the Glu/Asp-Arg-Tyr Motif of the alpha1b-Adrenergic Receptor: Effects on Receptor Isomerization and Activation. Molecular pharmacology 57: 219-231.
10. Scheer A, Fanelli F, Costa T, De Benedetti PG, Cotecchia S (1996) Constitutively active mutants of the alpha 1B-adrenergic receptor: role of highly conserved polar amino acids in receptor activation. EMBO J 15: 3566-3578.
11. Capra V, Veltri A, Foglia C, Crimaldi L, Habib A, et al. (2004) Mutational analysis of the highly conserved ERY motif of the thromboxane A2 receptor: alternative role in G protein-coupled receptor signaling. Molecular pharmacology 66: 880-889.
12. Shapiro DA, Kristiansen K, Weiner DM, Kroeze WK, Roth BL (2002) Evidence for a model of agonist-induced activation of 5-hydroxytryptamine 2A serotonin receptors that involves the disruption of a strong ionic interaction between helices 3 and 6. The Journal of Biological Chemistry 277: 11441-11449.

13. Greasley PJ, Fanelli F, Scheer A, Abuin L, Nenniger-Tosato M, et al. (2001) Mutational and Computational Analysis of the alpha1b-Adrenergic Receptor: Involvement of Basic and Hydrophobic Residues in Receptor Activation and G Protein Coupling. The Journal of Biological Chemistry 276: 46485-46494.
14. Zalokar M (1976) Autoradiographic Study of Protein and RNA Formation during Early Development of Drosophila Eggs. Dev Biol 49: 425-437.
15. Dahanukar A, Foster K, van der Goes van Naters WM, Carlson JR (2001) A Gr receptor is required for response to the sugar trehalose in taste neurons of Drosophila. Nat Neurosci 4: 1182-1186.
16. Kunwar PS, Sano H, Renault AD, Barbosa V, Fuse N, et al. (2008) Tre1 GPCR initiates germ cell transepithelial migration by regulating Drosophila melanogaster E-cadherin. J Cell Biol 183: 1-12.
17. Lerit DA, Gavis ER (2011) Transport of Germ Plasm on Astral Microtubules Directs Germ Cell Development in Drosophila. Current Biology 21: 439-448.
18. Fischer M, Haase I, Wiesner S, Müller-Taubenberger A (2006) Visualizing cytoskeleton dynamics in mammalian cells using a humanized variant of monomeric red fluorescent protein. FEBS Letters 580: 2495-2502.
19. Shcherbo D, Murphy CS, Ermakova GV, Solovieva EA, Chepurnykh TV, et al. (2009) Far-red fluorescent tags for protein imaging in living tissues. The Biochemical journal 418: 567-574.
20. Hanyu-Nakamura K, Kobayashi S, Nakamura A (2004) Germ cell-autonomous Wunen2 is required for germline development in Drosophila embryos. Development 131: 4545-4553.
21. Renault AD, Kunwar PS, Lehmann R (2010) Lipid phosphate phosphatase activity regulates dispersal and bilateral sorting of embryonic germ cells in Drosophila. Development 137: 1815-1823.
22. Sano H, Renault A, Lehmann R (2005) Control of lateral migration and germ cell elimination by the Drosophila melanogaster lipid phosphate phosphatases Wunen and Wunen 2. The Journal of Cell Biology 171: 675-683.
23. Starz-Gaiano M, Cho N, Forbes A, Lehmann R (2001) Spatially restricted activity of a Drosophila lipid phosphatase guides migrating germ cells. Development 128: 983-991.
24. Yamada Y, Davis KD, Coffman CR (2008) Programmed cell death of primordial germ cells in Drosophila is regulated by p53 and the Outsiders monocarboxylate transporter. Development 135: 207-216.
25. Mirzadegan T, Benkö G, Filipek S, Palczewski K (2003) Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. Biochemistry 42: 2759-2767.

APPENDIX

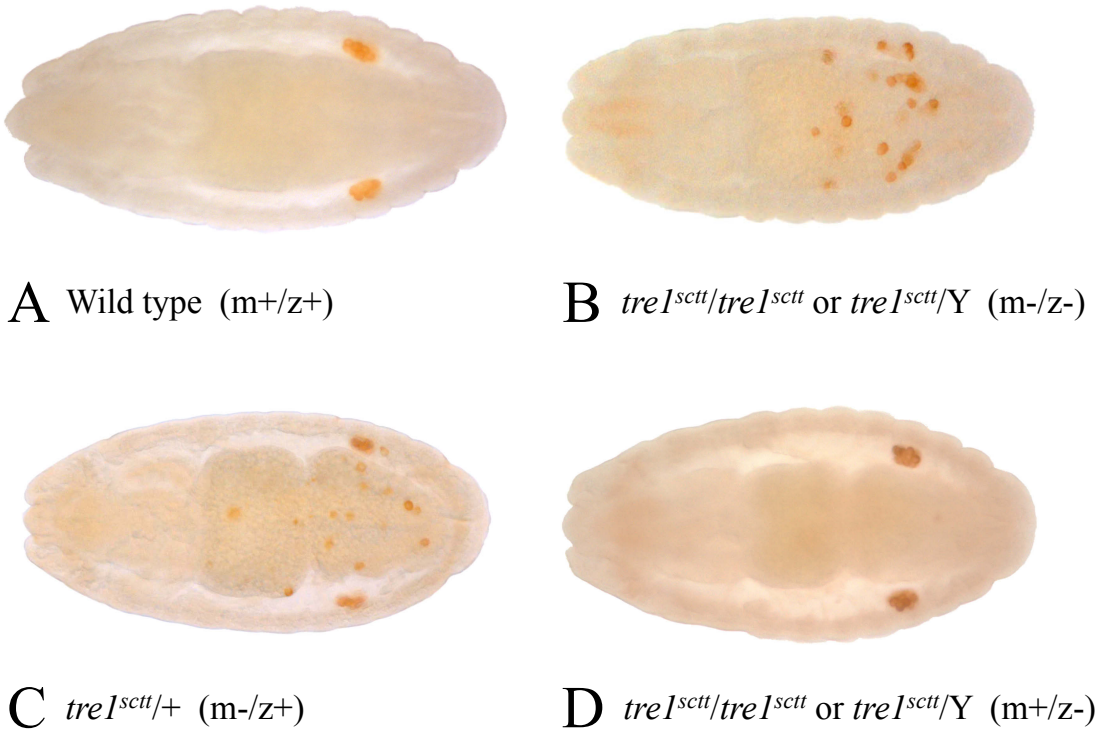
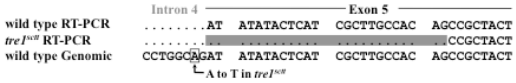
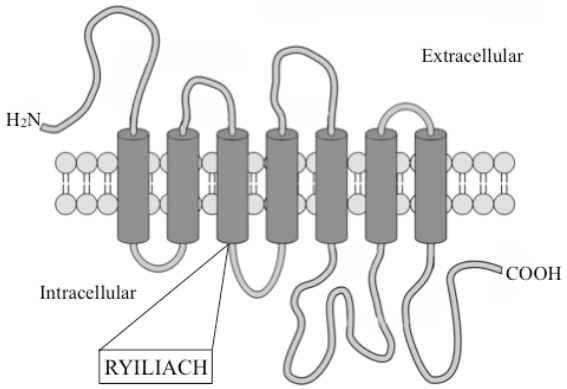


Figure 2.1. The *tre1*^{sctt} mutation disrupts germ cell migration. (A-D) Dorsal views of stage 15-16 embryos are shown. Anterior is to the left. Germ cells are labeled brown with an anti-Vasa antibody. (A) In wild-type embryos, the germ cells migrate to and coalesce with the somatic gonad precursor cells. (B) Germ cells do not migrate to the gonads in *tre1*^{sctt} m-/z- embryos. (C) Germ cell migration is restored in *tre1*^{sctt} m-/z+ embryos that have a wild-type *tre1* gene supplied paternally. (D) Germ cell migration is normal in *tre1*^{sctt} m+/z- embryos.

A



B



C

Receptor

Tre1	R	Y	I	L	I	A	C	H
GPR84	R	Y	L	L	I	A	-	H
CXCR4	R	Y	L	A	I	V	-	H
CXCR3	R	Y	L	N	I	V	-	H
PAR-2	R	Y	W	V	I	V	N	P
$S1P_3$	R	H	L	T	M	I	K	M
$S1P_4$	R	P	V	A	E	S	G	A
LPA_3	R	H	M	S	I	M	R	M

Figure 2.2. The *tre1*^{sctt} mutation results in an in-frame loss of eight amino acids from the third transmembrane domain/second intracellular loop junction of Tre1. (A) Reverse transcriptase PCR was performed on cDNA from 0-8 hour *tre1*^{sctt} and wild-type embryos. The *tre1*^{sctt} template reveals a deletion of 24 base pairs of exon 5, following the A to T base pair change in intron 4 of *tre1*^{sctt} mutants. The single base pair change is boxed. The nucleotides missing in *tre1*^{sctt} mRNA are highlighted. (B) Schematic diagram showing the predicted secondary structure of Tre1. SOSUI, TopPred, and TMHMM all predict the eight amino acid deletion results in the shortening of the second intracellular loop. However, the overall topology of the rest of the protein is unaffected. The missing amino acids, RYILIACH, are indicated. (C) A sequence alignment comparing the *tre1*^{sctt} amino acid deleted region to human GPCRs involved in cell migration. Identical residues are in dark gray and similar residues are in light gray.

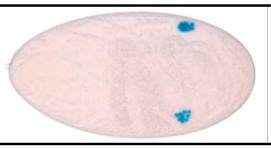
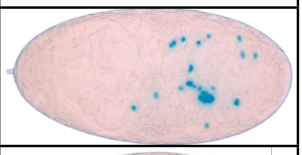
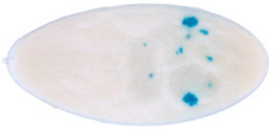
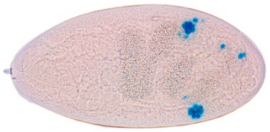
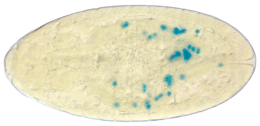
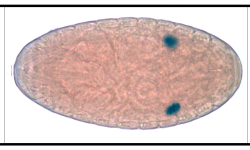
Rescues Migration		Does Not Rescue Migration	
Transgene	Phenotype	Transgene	Phenotype
RY ILI ACH		No transgene	
RY AAA AAA		*tre1*^{sctt} reconstruction	
RY AAA ACH		AY ILI ACH	
RA ILI ACH			

Figure 2.3. Transgenic rescue identifies the arginine of the E/N/DRY motif as a critical amino acid of Tre1 necessary for germ cell migration. Dorsal views of embryos are shown. Anterior is to the left. Stage 15-16 embryos were stained with X-Gal to visualize the *fat facets-lacZ* transgene, a germ cell marker. Embryos were from *tre1*^{sctt} homozygous mothers containing at least one copy of the specified transgene. The substituted amino acids are underlined. Replacement of the arginine with alanine results in a transgene that fails to rescue germ cell migration in *tre1*^{sctt} maternal-/zygotic- embryos. The *tre1*^{sctt} reconstruction is missing the 24 base pairs missing in *tre1*^{sctt} mutants.

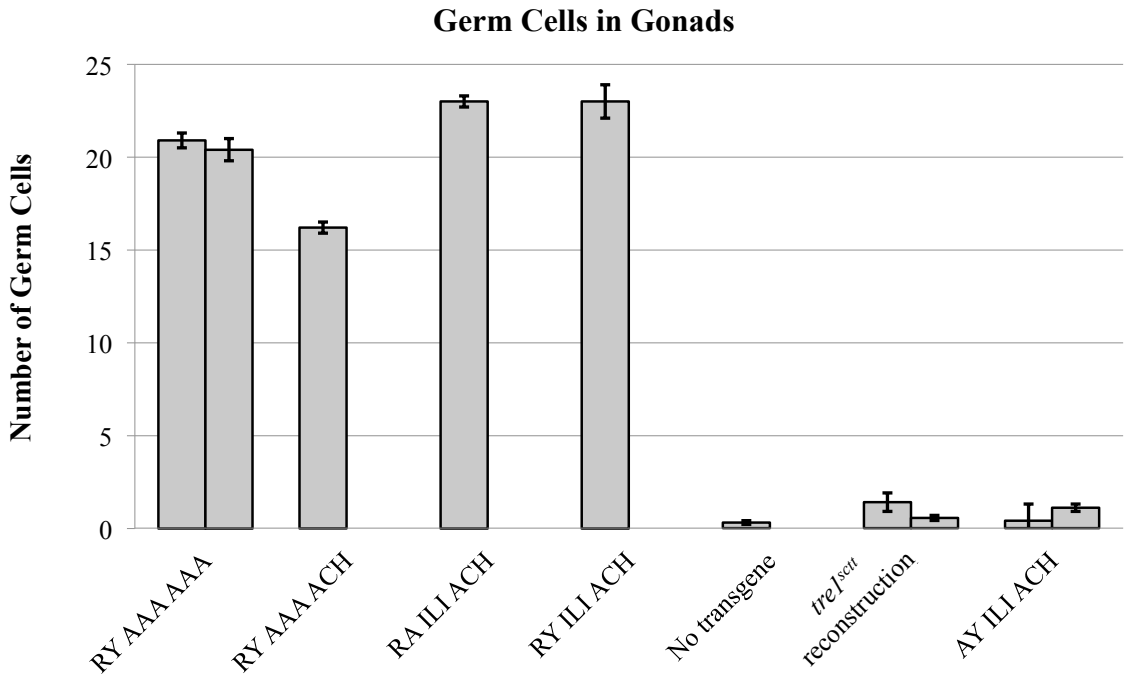


Figure 2.4. Germ cell counts showing the rescue of germ cell migration in *tre1*^{scti} mutants using transgenic constructs. The number of germ cells in the gonads of embryos in transgenic maternal rescue of the *tre1*^{scti} defect was analyzed. All test constructs assayed rescue germ cell migration with the exception of the arginine to alanine substitution, AYILIACH. The RYILIACH construct is the positive control and the no transgene and the *tre1*^{scti} reconstruction constructs are negative controls. Error bars represent the standard error of the mean (SEM).

Table 2.1. Germ cell distribution in *tre1*^{sctt} mutants.

Genotype			Germ Cells			
Maternal	**Paternal**	**Zygotic**	**In Gonads**	**Ectopic**	**Total**	**N**
wt/wt [a]	wt/Y	maternal+/zygotic+	14.7 ± 0.4	0.5 ± 0.1	15.2 ± 0.4	61
tre1^{sctt}/*tre1*^{sctt}	*tre1*^{sctt}/Y	maternal-/zygotic-	1.0 ± 0.2	16.3 ± 0.6	17.3 ± 0.7	69
tre1^{sctt}/*tre1*^{sctt}	wt/Y	maternal-/zygotic- [b]	0.3 ± 0.1	16.6 ± 0.8	16.9 ± 0.7	24
		maternal-/zygotic+ [b]	13.0 ± 0.5	7.0 ± 0.6	20.0 ± 0.8	23

Germ cell counts performed on 12-15 hour old embryos, Mean ± SEM. [a] wt denotes the non-mutagenized *w*^{1118}, *P{w*$^{+}$, *fat facets-lacZ}* parental strain. [b] Two distinct phenotypic classes, presumed genotypes are based on genotyping experiments performed in Coffman et al. 2002.

Table 2.2. Germ cell distribution in *tre1*^{sctt} maternal^{-}zygotic^{-} embryos from mothers with modified *tre1* transgenes.

Transgene	Germ Cells in Gonads	Germ Cells Ectopic	Total Number of Germ Cells	N
RY ILI ACH [a]	23.0 ± 0.9	1.2 ± 0.4	24.2 ± 0.7	69
RY AAA AAA [b]	20.9 ± 0.4	8.5 ± 0.4	29.4 ± 0.5	194
	20.4 ± 0.6	8.9 ± 0.7	29.3 ± 0.9	64
RY AAA ACH	16.2 ± 0.3	7.3 ± 0.8	23.5 ± 0.3	98
RA ILI ACH	23.0 ± 0.3	1.7 ± 0.3	24.7 ± 0.5	64
No transgene	0.3 ± 0.1	25.2 ± 0.9	25.5 ± 0.9	75
tre1^{sctt} reconstruction [b]	1.4 ± 0.5	23.3 ± 0.7	24.6 ± 0.7	103
	0.3 ± 0.2	22.1 ± 1.4	22.4 ± 1.4	29
AY ILI ACH [b]	0.4 ± 0.9	26.6 ± 0.8	27.1 ± 0.8	87
	1.1 ± 0.2	17.8 ± 0.9	18.9 ± 0.9	77

Embryos were collected and aged to stages 15-16, Mean ± SEM. Germ cells were detected by staining for β-galactosidase activity using the *P{w*$^{+}$*, fat facets-lacZ}* germ cell-specific marker. [a] Wild type T^{+}G^{+} vector as described in Dahanukar et al. 2001. [b] Two independent transgenic insertions were assayed.

Table 2.3. Synthetic oligonucleotides used for transgenic constructs to evaluate Tre1 function.

Transgene	Description of Amino Acid Change	Phosphorylated Oligos Used for Amino Acid Changes *
RY ILI ACH (T+G+ construct)	Rescue construct, wild type amino acids	
tre1^{sctt} reconstruction	The construct is missing the 24 base pairs missing from *tre1*^{sctt} RNA	5'-GTAGCGGCT----------GCCAGGGGATTA 5'-CCCTGGCAG----------CCGCTACTCGCA
RY AAA AAA	ILI ACH replaced with 6 alanines	5'-GTAGCGGCTGGCGGCGGCGGCGGCATATCTGCCAGGGGATTA 5'-CCCTGGCAGATATGCCGCCGCCGCCGCCGCCAGCCGCTACTCGCA
RY AAA ACH	ILI replaced with 3 alanines	5'-GTAGCGGCTGTGGCAAGCGGCGGCGGCATATCTGCCAGGGGATTA 5'-CCCTGGCAGATATGCCGCCGCCGCTTGCCACAGCCGCTACTCGCA
RA ILI ACH	Tyrosine replaced with an alanine	5'-GTAGCGGCTGTGGCAAGCGATGAGTATGGCTCTGCCAGGGGATTA 5'-CCCTGGCAGAGCCATACTCATCGCTTGCCACAGCCGCTACTCGCA
AY ILI ACH	Arginine replaced with an alanine	5'-GGCCTACGCGTTCAGGGTGATGCCCACCATGC 5'-GTGGGCATCACCCTGAACGCGTAGGCCCAAGT

* Underlined sequence designate the nucleotide replacements used to create the amino acid substitutions

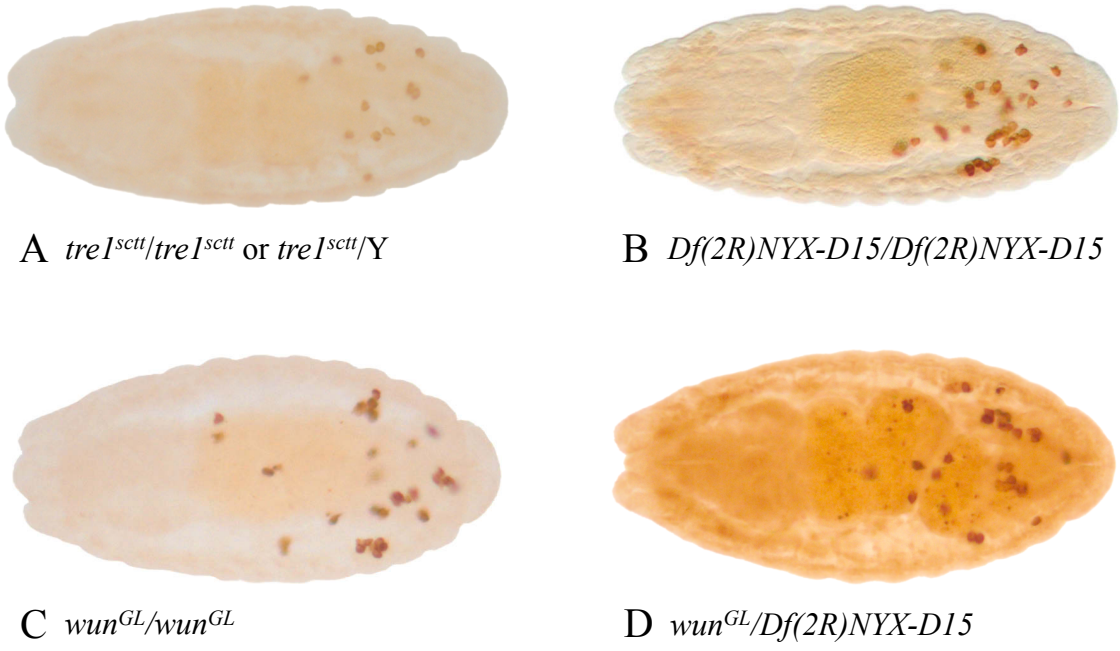


Figure 3.1. *wun wun2* mutant embryos have severe germ cell phenotypes similar to *tre1*^{sctt} mutant embryos. Dorsal views of stages 15-16 embryos are shown. Anterior is to the left. The primordial germ cells are labeled using a chicken anti-Vasa antibody. Homozygous *wun wun2* embryos are shown and were identified by the absence of anti-β-galactosidase staining of the balancer CyO, P[*ry*$^{+}$*, ftz-lacZ*]. (A) A *tre1*^{sctt}/*tre1*^{sctt} or *tre1*^{sctt}/*Y* embryo. (B) A *Df(2R)NYX-D15/Df(2R)NYX-D15* embryo. (C) A *wun*^{GL}/*wun*^{GL} embryo. (D) A *wun*^{GL}*/Df(2R)NYX-D15* embryo. This embryo was the result of crossing a *Df(2R)NYX-D15*/CyO, P[*ry*$^{+}$*, ftz-lacZ*] female to a *wun*^{GL}/CyO, P[*ry*$^{+}$*, ftz-lacZ*] male. In all embryos, primordial germ cell migration is severely disrupted with few primordial germ cells reaching the gonads (Table 3.2).

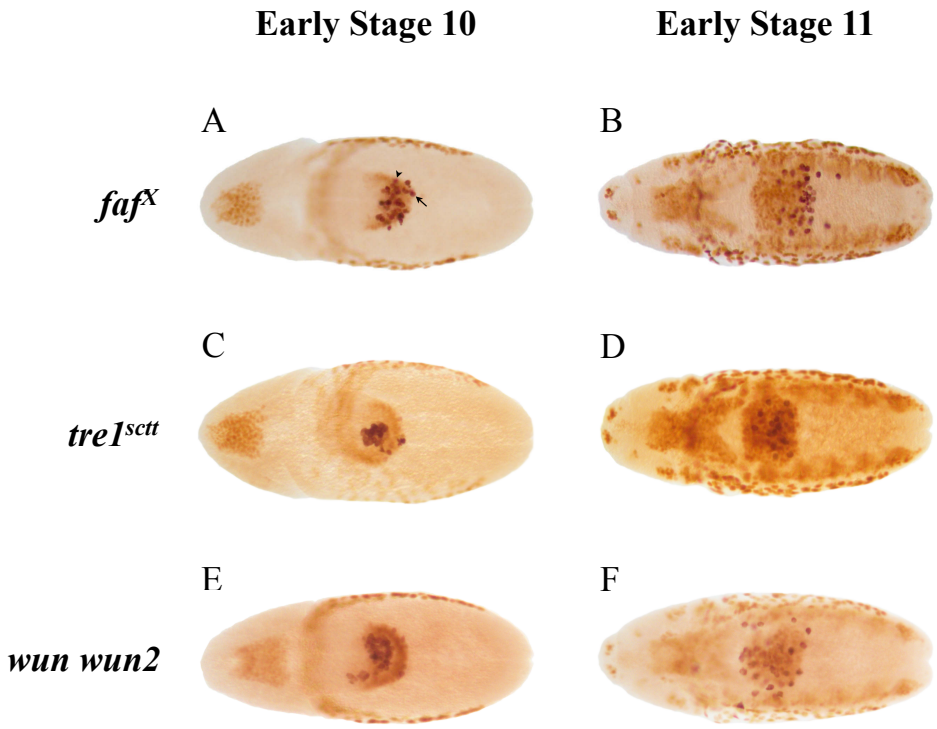


Figure 3.2. Analyzing primordial germ cell migration across the posterior midgut epithelium during stages 10 and 11. Dorsal views are shown and anterior is to the left. The primordial germ cells are labeled purple using a chicken anti-Vasa antibody and the VIP substrate. The anterior and posterior midgut, the amnioserosa, and parts of the nervous system are labeled brown using an anti-Hindsight antibody and the DAB substrate. (A, B) w^{1118}, P{w^{+}, *fat facets-lacZ*} (*faf* ^{X}) embryos. *faf* ^{X} is used as a control. (C, D) *tre1*^{sctt} embryos. (E, F) *wun wun2* embryos (*Df(2R)NYX-D15/Df(2R)NYX-D15*). (A, C, E) Early stage 10 embryos. At this stage, control and *wun wun2* primordial germ cells have already begun active migration. Additional primordial germ cells that are crossing the posterior midgut are above and below the focal plane shown in (E). Germ cell counts were performed on these embryos and the primordial germ cells were placed into two categories (Table 3.3). A primordial germ cell was scored as "crossing" (designated by an arrowhead) if the primordial germ cell was on the same focal plane as the Hindsight-stained midgut and near the periphery of the staining. A primordial germ cell was scored as "exited" (indicated by an arrow) if the primordial germ cell was outside of all Hindsight staining. (B, D, F) Early stage 11 embryos. At this stage, most of the primordial germ cells have exited the midgut in control and *wun wun2* embryos, but not in *tre1*^{sctt} embryos.

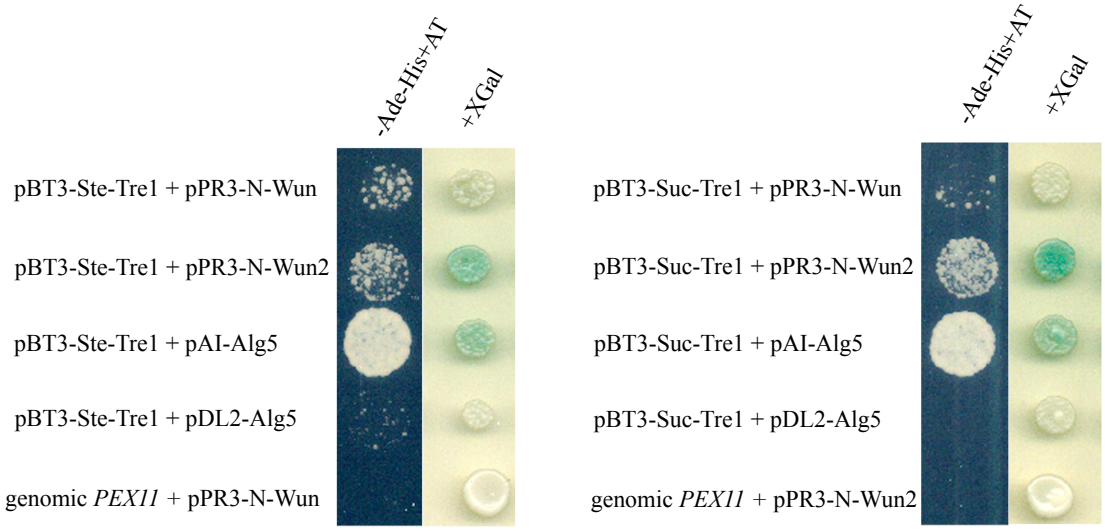


Figure 3.3. Tre1 interacts with Wun and Wun2 in a yeast two-hybrid assay. Yeast cells containing the *tre1* bait plasmids, pBT3-*STE-tre1* and pBT3-*SUC-tre1*, were transformed with *wun* and *wun2* prey plasmids (pPR3-N-*wun* and pPR3-N-*wun2*). Transformants were spotted on plates lacking adenine and histidine and supplemented with 5 mM aminotriazole (AT). These plates show that Tre1 interacts with both Wun and Wun2, although the growth with Wun2 is greater than with Wun. The strength of the interactions was further tested using a β-galactosidase assay. Similar to the colony growth assay, the Tre1-Wun2 interaction is stronger than the Tre1-Wun interaction with both pBT3-Ste and pBT3-Suc baits. pAI-Alg5 is the positive control prey and pDL2-Alg5 is the negative control prey. *PEX11* is a non-interacting bait construct that, when transformed with pPR3-N-*wun* or pPR3-N-*wun2*, shows no interaction, validating the specificity of this assay.

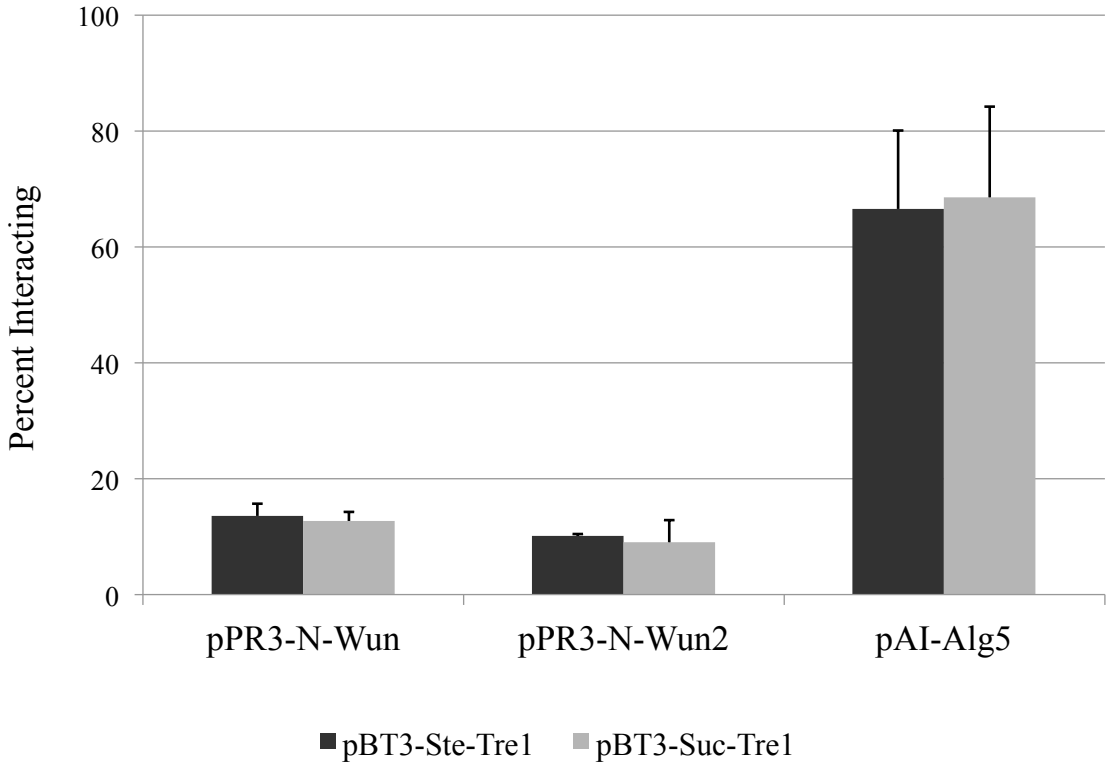


Figure 3.4. Colony growth assay indicates an interaction between Tre1 and Wun and Tre1 and Wun2. Yeast cells transformed with a bait vector (pBT3-*STE-tre1* or pBT3-*SUC-tre1*) and a prey vector (pPR3-N-*wun* or pPR3-N-*wun2*) were grown on selective (-Ade -His) and nonselective media. The percent interacting is the number of colonies on selective medium divided by the number of colonies on nonselective medium. Three independent experiments were performed and the average percent interacting for the three experiments is shown. Error bars represent standard error of the mean. For the pPR3-N-Wun prey, an average of 13.6% and 12.7% of colonies grew with the pBT3-Ste-Tre1 and pBT3-Suc-Tre1 baits, respectively. For the pPR3-N-Wun2 prey, an average of 10.1% and 9.0% of colonies grew with the pBT3-Ste-Tre1 and pBT3-Suc-Tre1 baits, respectively. For pAI-Alg5, the positive control bait, an average of 66.6% and 68.6% of colonies grew with pBT3-Ste-Tre1 and pBT3-Suc-Tre1 baits, respectively. The colony counts were normalized by subtracting the number of colonies observed with pDL2-Alg5, a negative control prey.

Table 3.1. Genes tested in a loss-of-function screen for disruptions in primordial germ cell development.

Gene targeted	Allele	Gene Class	Phenotype
crk	KG00336	SH2/SH3	Normal
gprk2	6936	Kinase	Normal
	EY09213	Kinase	Normal
jafrac2	f01922	Thioredoxin peroxidase	Normal
kurtz	c01503	Arrestin	Normal
neur	KG06174	E3 ubiquitin ligase	Normal
	11	E3 ubiquitin ligase	Normal
	A101	E3 ubiquitin ligase	Normal
pxt	EY03052	Peroxidase	Normal
src64B	KG00213	Src kinase	Normal
	EY16432	Src kinase	Normal
tao-1	EP1455	Kinase	Normal
	e04532	Kinase	Normal
trio	6A	Rho GEF	Normal
	KG06642	Rho GEF	Normal
wun	k10201	Lipid phosphate phosphatase	Germ cell migration/germ cell death
wun and *wun2*	GL	Lipid phosphate phosphatase	Germ cell migration/germ cell death
	CE	Lipid phosphate phosphatase	Germ cell migration/germ cell death
	NYX-D15	Lipid phosphate phosphatase	Germ cell migration/germ cell death
	GL/NYX-D15*	Lipid phosphate phosphatase	Germ cell migration/germ cell death
wun2	EP2650ex34	Lipid phosphate phosphatase	Germ cell migration/germ cell death

*GL/NYX-D15 is the result of crossing wun^{GL}/CyO, P[ry^{+}, *ftz-lacZ*] to Df(2R)*NYX-D15*/CyO, P[ry^{+}, *ftz-lacZ*] to get wun^{GL}/Df(2R)*NYX-D15* embryos.

Table 3.2. *wun wun2* and *tre1*^{sctt} mutant embryos have similar germ cell phenotypes at stages 15-16.

Embryo Genotype	Genotype of Parents, Female x Male	Germ Cells In Gonads		Ectopic Germ Cells		N
		Average ± SEM	Range	Average ± SEM	Range	
faf^{X}/*faf*^{X} or *faf*^{X}/*Y*	*faf*^{X}/*faf*^{X} × *faf*^{X}/*Y*	15.5 ± 0.5	10 – 20	1.1 ± 0.2	0 – 3	20
tre1^{sctt}/*tre1*^{sctt} or *tre1*^{sctt}/*Y*	*tre1*^{sctt}/*tre1*^{sctt} × *tre1*^{sctt}/*Y*	1.0 ± 0.2	0 – 2	18.6 ± 1.4	7 – 31	20
Df(2R)NYX-D15/Df(2R)NYX-D15	*Df(2R)NYX-D15/CyO* × *Df(2R)NYX-D15/CyO*	2.2 ± 0.3	0 – 9	34.7 ± 1.0	22 – 50	50
wun^{GL}/*wun*^{GL}	*wun*^{GL}/*CyO* × *wun*^{GL}/*CyO*	4.7 ± 0.4	1 – 10	26.2 ± 0.9	15 – 39	50
wun^{GL}/*Df(2R)NYX-D15*	*wun*^{GL}/*CyO* × *Df(2R)NYX-D15/CyO*	4.6 ± 0.5	0 – 9	29.4 ± 0.8	21 – 38	25
wun^{GL}/*Df(2R)NYX-D15*	*Df(2R)NYX-D15/CyO* × *wun*^{GL}/*CyO*	3.2 ± 0.4	0 - 11	29.7 ± 0.9	15 - 40	25

Stages 15-16 embryos were used for germ cell counts. *w*^{1118}, P{*w*$^{+}$, *fat facets-lacZ*} (*faf*^{X}) flies were used as the control. The *tre1*^{sctt} allele is recessive and shows a maternal effect (Coffman et al., 2002). All CyO balancers were CyO, P[*ry*$^{+}$, *ftz-lacZ*]. Homozygous embryos were distinguished by a lack of β-galactosidase staining of the CyO, P[*ry*$^{+}$, *ftz-lacZ*] balancer. All lines were stained with anti-EYA to confirm the somatic gonadal precursor cells were present (data not shown). SEM = standard error of the mean, N = total embryos scored.

Table 3.3. The initiation of primordial germ cell migration across the posterior midgut epithelium is delayed in *tre1*^{sctt} embryos, but not in *wun wun2* embryos.

Embryo Genotype	Early Stage 10					Late Stage 10 – Early Stage 11				
	Crossing	Range	Exited	Range	N	Crossing	Range	Exited	Range	N
faf^{X}/*faf*^{X} or *faf*^{X}/Y	5.6 ± 0.5	2 – 12	16.8 ± 1.5	3 – 28	20	0.1 ± 0.1	0 – 2	34.6 ± 0.8	29 – 43	20
tre1^{sctt}/*tre1*^{sctt} or *tre1*^{sctt}/*Y*	2.1 ± 0.3	0 – 5	1.9 ± 0.3	0 – 5	20	3.8 ± 0.4	0 – 8	5.1 ± 0.7	0 – 12	20
Df(2R)NYX-D15/Df(2R)NYX-D15	5.4 ± 0.6	0 – 11	12.8 ± 1.3	3 – 24	20	7.1 ± 0.9	0 – 14	28.4 ± 1.9	12 – 45	20

w^{1118}, P{w^{+}, *fat facets-lacZ*} (*faf*^{X}) flies were used as the control. Crossing indicates the average number of primordial germ cells crossing the posterior midgut ± the standard error of the mean (SEM). Exited indicates the average number of primordial germ cells that have exited the posterior midgut ± SEM. The primordial germ cells in the posterior midgut were not scored. N = total embryos scored.

Table 3.4. *faf*^{X}, *tre1*^{sctt} and *wun wun2* mutant embryos have similar numbers of primordial germ cells at stages 5-6.

Embryo Genotype	Total Number of Germ Cells		N
	Average ± SEM	Range	
faf^{X}/*faf*^{X} or *faf*^{X}/*Y*	35.3 ± 2.5	29 – 45	6
tre1^{sctt}/*tre1*^{sctt} or *tre1*^{sctt}/*Y*	34.5 ± 2.4	26 – 43	6
Df(2R)NYX-D15/Df(2R)NYX-D15	36.5 ± 2.9	22 – 41	6
wun^{GL}/*wun*^{GL}	37.7 ± 2.1	32 – 45	6

Stages 5-6 embryos were used for germ cell counts. *w*^{1118}, P{*w*$^{+}$, *fat facets-lacZ*} (*faf*^{X}) flies were used as the control. Homozygous *wun wun2* embryos were distinguished by a lack of β-galactosidase staining of the CyO, P[*ry*$^{+}$, *ftz-lacZ*] balancer. SEM = standard error of the mean, N = total embryos scored.

Table 3.5. Constructs tested in the split-ubiquitin yeast two-hybrid screen when Tre1 was the bait.

Prey	Interaction?
Crk	No
Csk	No
Dock	No
GPRK1	No
GPRK2	No
iHog	No
Jafrac1	No
Kurtz	No
Neur	No
Patched	No
Ptp61F	No
Rho1	No
Smoothened	No
Tao	No
Wun	Yes
Wun2	Yes

The baits used to test all preys were pBT3-Ste-Tre1 and pBT3-Suc-Tre1. If no significant growth was observed, the prey was marked as "No Interaction."

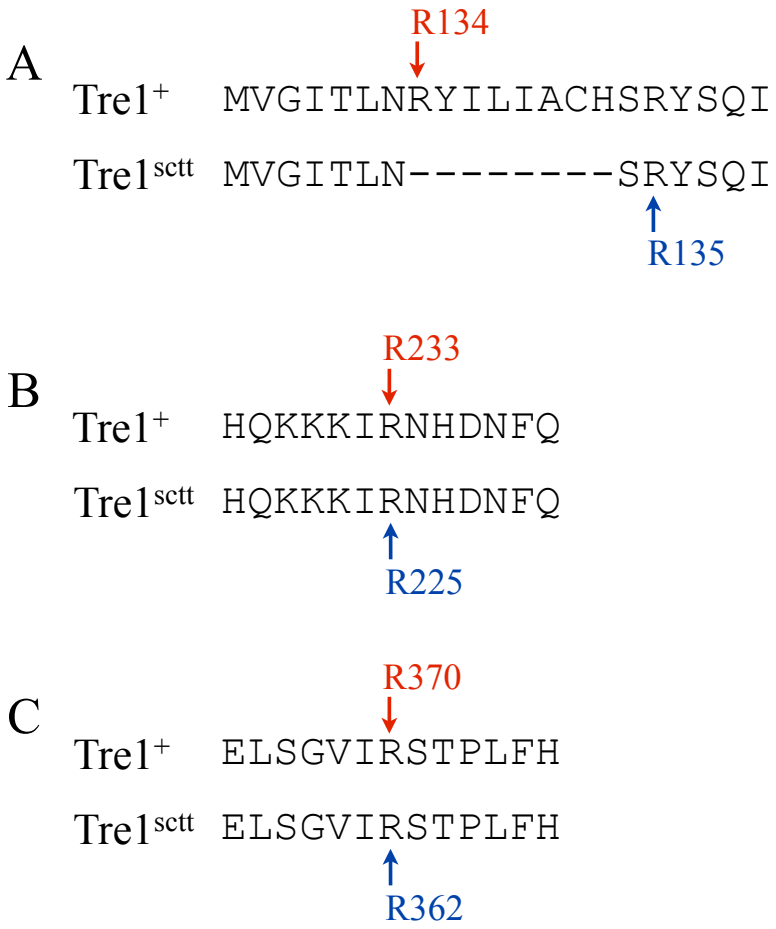


Figure 4.1. Amino acid numbering scheme. Amino acid residues are labeled using the single-letter code followed by the amino acid absolute sequence number. (A) Some of the amino acids in the transmembrane 3 – intracellular loop 2 region. The arginine of interest in this region is R134 in Tre1^{+}. The corresponding arginine in Tre1^{sctt} is missing, but an alternative arginine is studied, R135. (B) Some of the amino acids of intracellular loop 3. R233 in Tre1^{+} and R225 in Tre1^{sctt} are used as controls in the solvent accessible surface area (SASA) calculation. (C) Some of the amino acids in the C-terminal tail. Like the arginines in intracellular loop 3, R370 in Tre1^{+} and R362 in Tre1^{sctt} are used as controls in the SASA calculation.

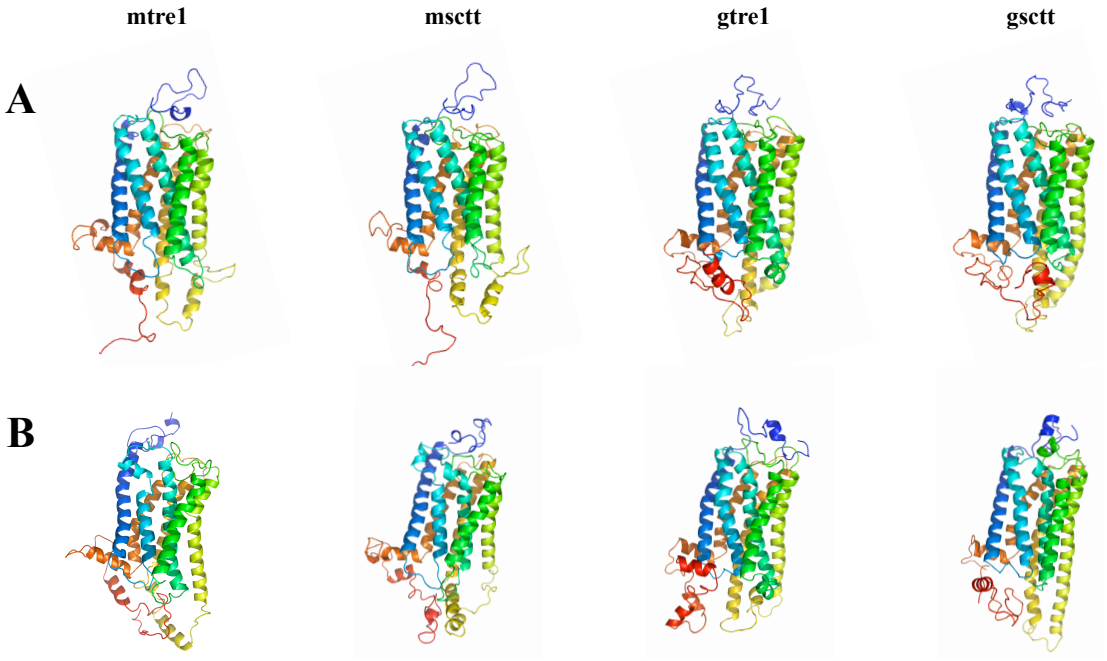


Figure 4.2. Three-dimensional models for Tre1^{+} and Tre1^{sctt}. mtre1 and msctt are the models generated with GPCR-ModSim and gtre1 and gsctt are the models generated with GPCR-ITASSER. mtre1 and gtre1 are models for Tre1^{+} and msctt and gsctt are models for Tre1^{sctt}. The N-termini are colored blue and the C-termini are colored red. (A) The best models chosen from the GPCR-ModSim and GPCR-ITASSER. (B) The resulting models after 200 ns of molecular dynamics (MD) simulations when the protein was embedded in a POPE lipid bilayer (lipids, water and ions are not shown). Some of the differences between the structures generated by GPCR-ModSim and GPCR-ITASSER are still present after 200 ns of MD. For example, in the models from GPCR-ModSim, helixes 5 and 6 (yellow and gold) are similar in length to the other five helices. Helixes 5 and 6 from GPCR-ITASSER are longer than the other helices.

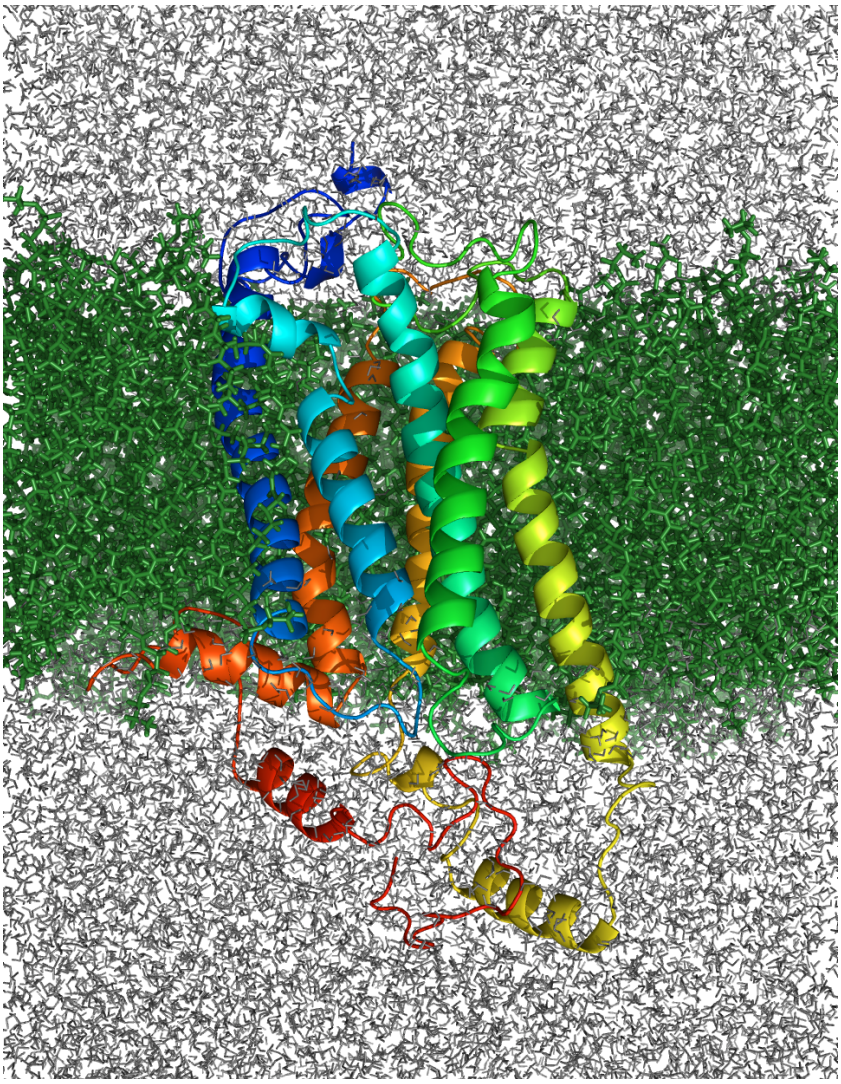


Figure 4.3. An example of a protein structure prediction of Tre1 embedded in a POPE lipid bilayer. The protein structure prediction modeled here is mtre1 after 200 ns of molecular dynamics. The extracellular surface of the bilayer is on top and the intracellular surface of the bilayer is on the bottom. mtre1 is depicted as ribbons, with the N-terminus colored blue and the C-terminus colored red. The POPE bilayer is represented as sticks and is colored green. Water and ions are depicted as grey lines.

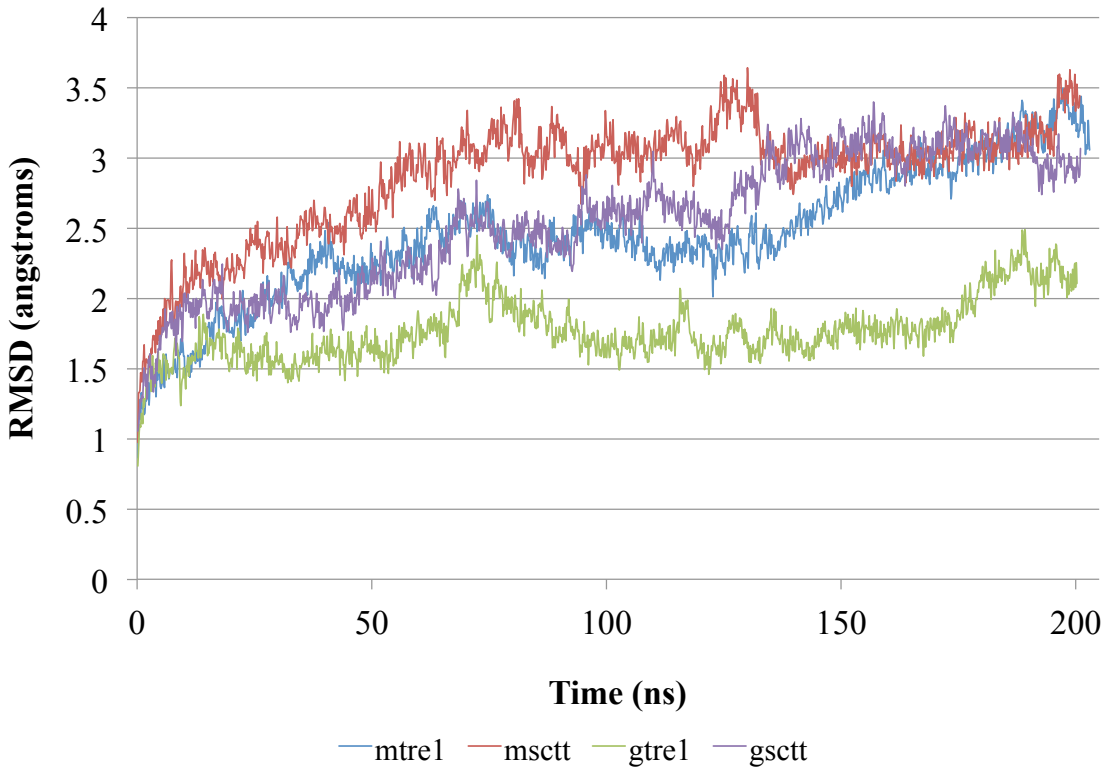


Figure 4.4. Root mean squared deviation of the transmembrane regions of the protein models versus time. Root mean squared deviation (RMSD) was calculated for the Cα atoms of the transmembrane domains and is plotted over simulation time. The transmembrane regions of the protein models do not change much compared to the starting structures over the course of 200 ns of molecular dynamics, as the final RMSD values are between 2 – 3.5 Å.

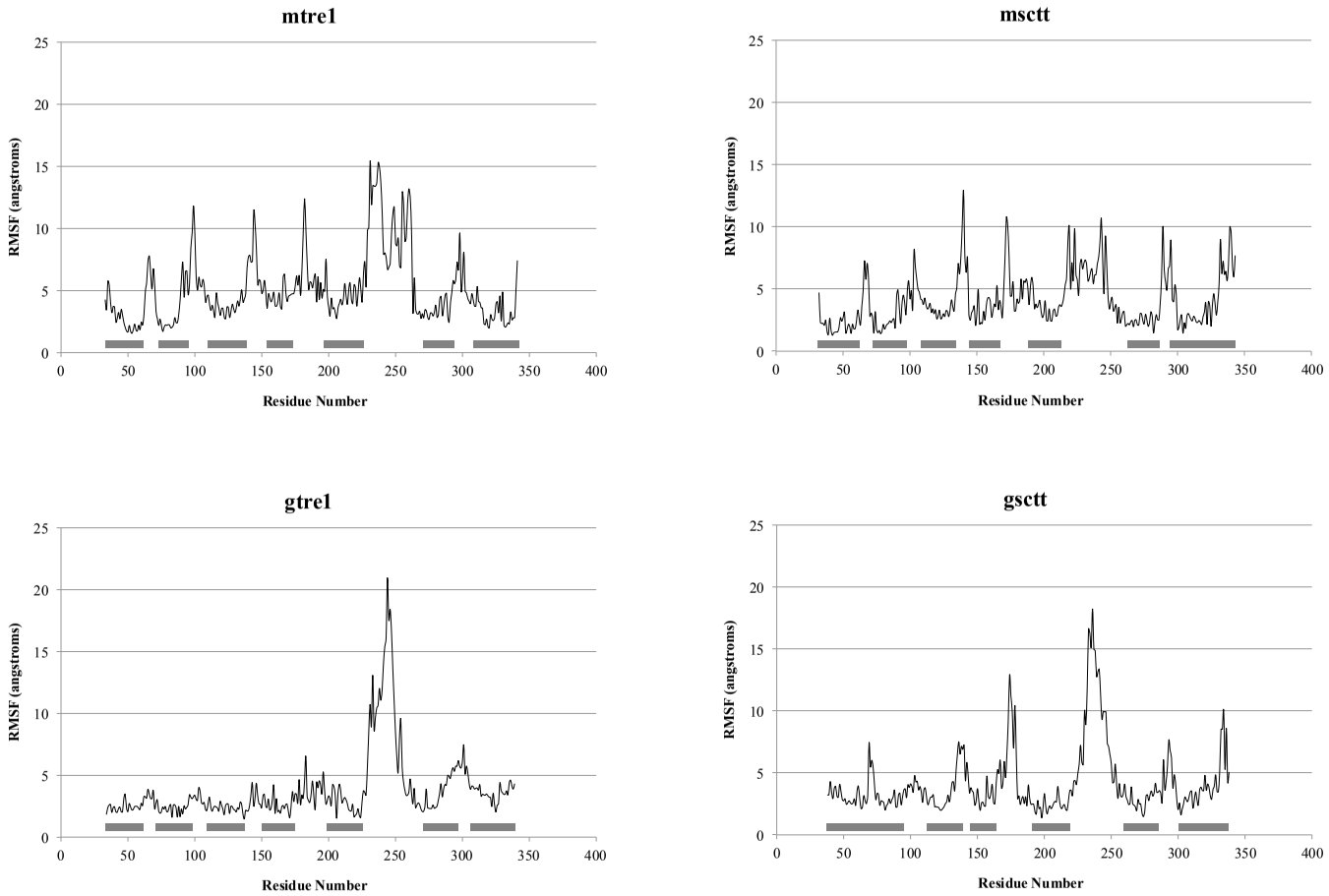


Figure 4.5. Root mean squared fluctuation of the mtre1, msctt, gtre1 and gsctt model systems. The root mean squared fluctuation (RMSF) for each model system in this study are shown. The N- and C-termini are not included in the plots since the termini had high RMSF values and made it difficult to see the fluctuations in the other regions of the protein. The dark grey boxes denote the regions of the protein that were within the lipid bilayer. The regions of the protein within the lipid bilayer had much lower RMSF values than the loop regions.

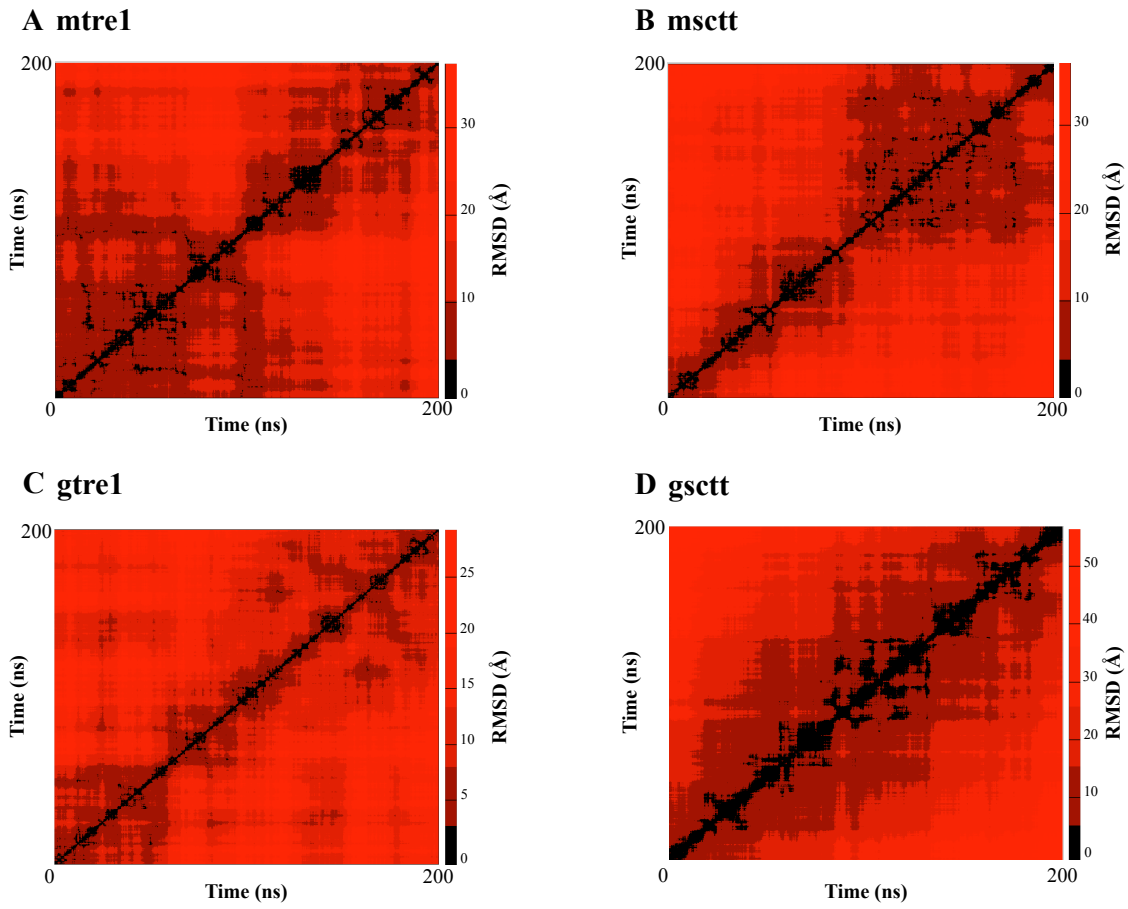


Figure 4.6. Heatmaps of an all-to-all RMSD calculation on transmembrane Cα atoms in mtre1, msctt, gtre1 and gsctt model systems. An all-to-all RMSD calculation is when the RMSD of all pairs of frames in a trajectory are computed. This all-to-all RMSD calculation was performed using the VMD plug-in RMSD Trajectory Tool (Humphry et al., 1996). The RMSD scale is different for each model system and is located to the right of each plot. The diagonal block structures in each plot show the Tre1 protein ($Tre1^{+}$ or $Tre1^{sctt}$) exploring the conformational space. As no diagonal block structure was revisited, it suggests a lack of convergence.

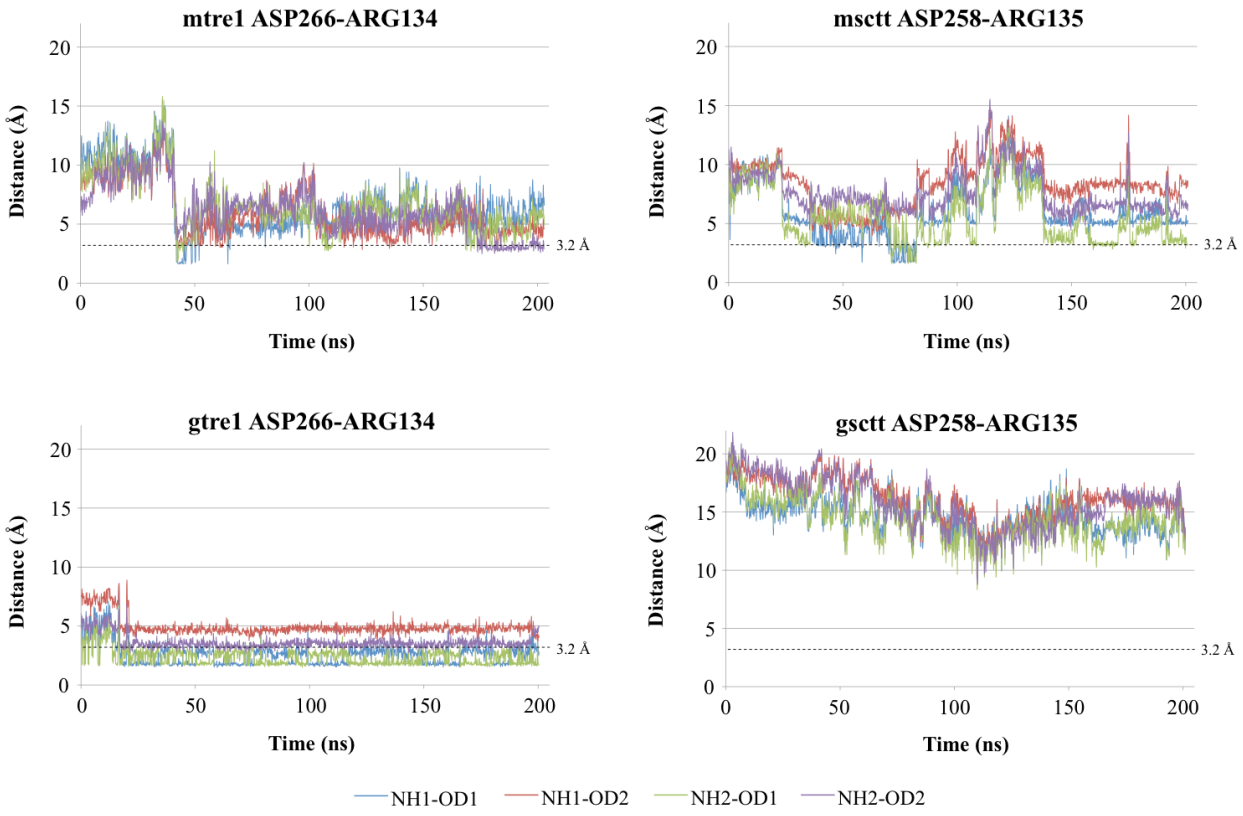


Figure 4.7. Distances between an aspartic acid of transmembrane helix 6 and an arginine of transmembrane helix 3 are close enough to form a salt bridge in mtre1, msctt and gtre1 model systems. Distance between oxygens of the carboxylate group of aspartic acid (ASP) (OD1, OD2) and nitrogens of the guanidium group of arginine (ARG) (NH1, NH2) are plotted. N-O distances of 3.2 Å or less are capable of forming a salt bridge. The N-O distances are close enough to form a salt bridge transiently in mtre1 and msctt. There is potential for a salt bridge to be consistently present in the gtre1 model system. The N-O distances are too great in gsctt to form a salt bridge at any point during the simulation.

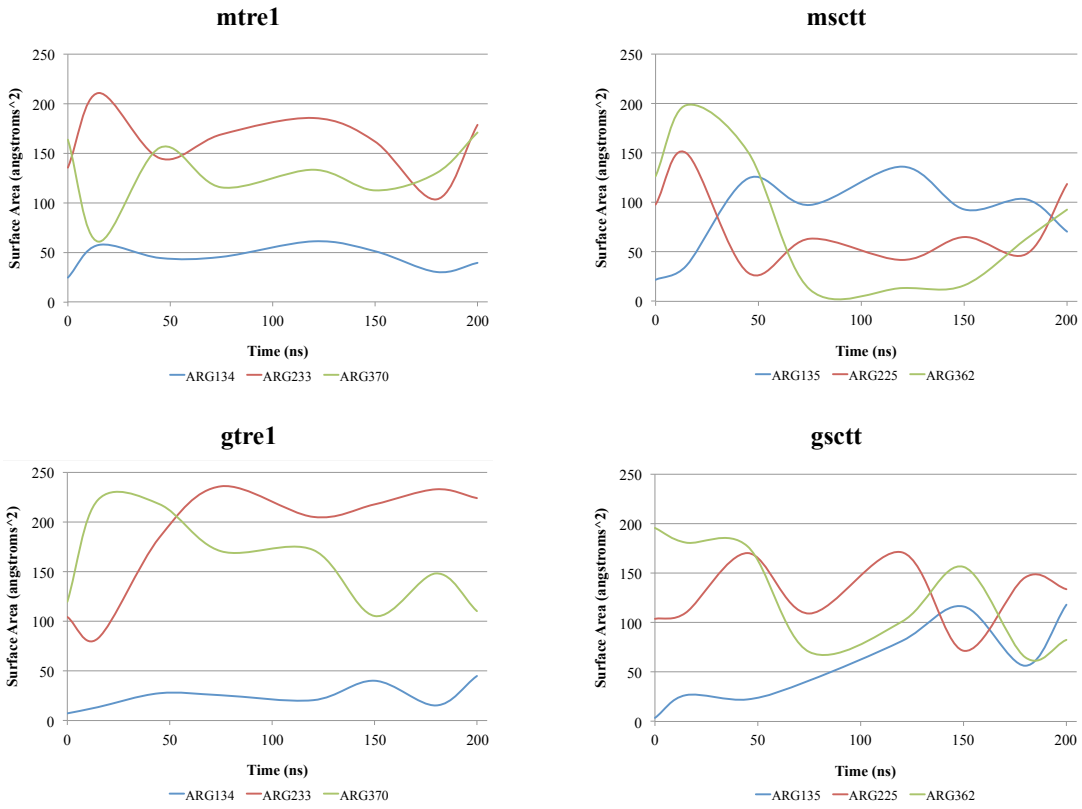


Figure 4.8. Solvent accessible surface area of selected arginine residues. The solvent accessible surface area (SASA) was calculated eight times during the trajectories for the different model systems. SASA was calculated for the arginine of the NRY motif (ARG134) in mtre1 and gtre1. ARG135 was used in this calculation for the msctt and gsctt systems as a potential alternative arginine for salt bridge formation. Control arginines are located in intracellular loop 3 (ARG223 or ARG225) or the C-terminal tail (R370 or R362).

Table 4.1. Multiple sequence alignment and template selection in GPCR-ModSim

Template	Tre1^{+} % Identity	Tre1^{sctt} % Identity
1U19 – bovine rhodopsin	14.3	14.1
2RH1 – human β_2-adrenergic receptor	15.1	14.6
2VT4 – turkey β_1-adrenergic receptor	14.8	14.3
2Z73 – squid rhodopsin *	17.4	16.7
3EML – human A_{2A}-adenosine receptor	14.0	13.5
3ODU – human chemokine receptor 4	16.3	14.8
3PBL – human D_3 dopamine receptor	16.1	15.9

Results were obtained from GPCR-ModSim (Rodriguez et al., 2012). * Denotes the template chosen for homology modeling.

Table 4.2. Voronoi tessellation Monte Carlo integration method confirms the model systems have maintained a fluid-phase bilayer.

Model System	Average Surface Area (Å^2/lipid)	
	Boundary Lipid	Non-Boundary Lipid
mtre1	47.4	59.9
msctt	44.2	61.0
gtre1	45.3	59.8
gsctt	47.0	60.9

Surface areas were calculated using the Voronoi tessellation Monte Carlo integration method (Mori et al., 2012) and calculated at 200 ns for all model systems. A boundary lipid is a lipid that contacts an atom from the protein and a non-boundary lipid is a lipid that makes no contact with protein atoms. Experimental surface area per non-boundary POPE lipid is 59.75 – 60.75 Å^2 (Rappolt et al., 2003). Computationally determined surface area per non-boundary POPE lipid using CHARMM36 is 59.2 Å^2 (Klauda et al., 2010).

Table 4.3. Approximate decorrelation times for the different model systems.

Model System	τ_d1 [a] (ns)	τ_d1 [b] (ns)	τ_d2 [a] (ns)	τ_d2 [b] (ns)
mtre1	15.3 – 21.3	11.7 – 15.9	119.5 – 141.2	56.7 – 63.5
msctt	18 – 23.4	10.8 – 14.4	125.2 – 138.5	57.6 – 65.6
gtre1	22.5 *	14.1 – 15	105.6 – 132.9	50.6 – 60.7
gsctt	13.8 – 18.3	13.5 – 14.4	123.6 – 140.1	61.7 – 68.6

τ_d1 is the decorrelation time as estimated from the plot of σ^2 (t) with step sizes 2, 4 and 5. The decorrelation time calculation is from Lyman and Zuckerman, 2007. τ_d2 is the decorrelation time from the automated effective sample size calculation, developed by Zhang et al., 2010. Both calculations are part of the LOOS analysis library (Romo and Grossfield, 2009). [a] 10 bins (reference structures) were used, [b] 20 bins were used. * The decorrelation time for the gtre1 system with 10 bins is estimated from the step size of 2 curve only.

Table 4.4. Assessing convergence of the different model systems using the blocked covariance overlap method.

Model System	k_1	t_1 (ns)	k_2	t_2 (ns)	k_3	t_3 (ns)	BBCOM/BCOM
mtre1	4.5	0.9	2.0	4.9	2.0	83.0	1.5
msctt	2.8	0.9	1.6	4.7	1.8	75.8	1.5
gtre1	2.0	0.8	0.8	7.1	0.8	128.2	1.3
gsctt	3.9	0.8	1.7	5.2	1.7	81.8	1.6

BBCOM is the bootstrapped blocked covariance overlap method and BCOM is the blocked covariance overlap method. t_1 - t_3 are decorrelation times from fitting the BBCOM/BCOM curve to: $f(t) = k_1e^{-t/t1} + k_2e^{-t/t2} + k_3e^{-t/t3} + 1$ (Romo and Grossfield, 2011). The BBCOM/BCOM ratio decays to a final ratio of greater than 1 for each model system. This suggests that the systems have not yet converged. BBCOM/BCOM is part of the LOOS analysis library (Romo and Grossfield, 2009).

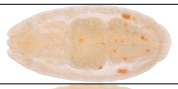
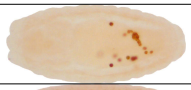
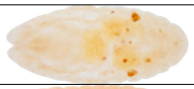
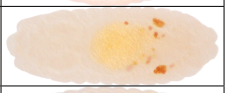
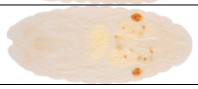
Rescues Migration		Does Not Rescue Migration	
Male	Phenotype	Male	Phenotype
faf^{X}		No transgene	
RY ILI ACH		*tre1*^{sctt} reconstruction	
RA ILI ACH		AY ILI ACH	
RY AAA ACH		RY AAA ACH	
RY AAA AAA		RY AAA AAA	
RY AAA AAA recomb (1)			
RY AAA AAA recomb (2)			
RY AAA AAA recomb (3)			

Figure 5.1. Zygotic expression of *tre1* transgenes reveals a role for the amino acids ILIACH. Dorsal views of stages 15-16 embryos are shown. Anterior is to the left. The primordial germ cells are labeled using a chicken anti-Vasa antibody. Embryos are from a homozygous *tre1*^{sctt} female crossed to a *tre1*^{sctt} male carrying the specified transgene. When control (*faf*^{X}), RYILIACH or RAILIACH males are used, germ cell migration is rescued, as determined by incorporation of more than 12 primordial germ cells into the gonads (Table 5.1). Conversely, when no transgene, *tre1*^{sctt} reconstruction or AYILIACH males are used, there is no rescue of migration. The inability of the arginine substitution (AYILIACH) to rescue migration, in agreement with the maternal rescue experiments, demonstrates a critical role for the arginine in Tre1 function. The middle panel shows results for the two constructs RYAAAACH and RYAAAAAA. When RYAAAACH or RYAAAAAA are supplied paternally, both partial rescue of migration and no rescue of migration phenotypes are observed. To determine if the partial rescue of migration seen with RYAAAAAA was due to a dosage effect, recombinant lines were constructed that contained two copies of RYAAAAAA on each second chromosome. The bottom panel pictures results from the recombinant lines and shows that when the dosage of RYAAAAAA is increased, primordial germ cell migration is more completely rescued.

Table 5.1. Germ cell distribution when *tre1* transgenes are paternally supplied.

Male	Germ Cells In Gonads		Ectopic Germ Cells		Total Germ Cells		N
	Average ± SEM	Range	Average ± SEM	Range	Average ± SEM	Range	
RY ILI ACH	16.2 ± 0.5	9 - 27	9.7 ± 0.5	3 - 19	25.9 ± 0.7	10 - 38	25
no transgene	1.2 ± 0.4	0 - 7	15.4 ± 1.1	7 - 25	16.5 ± 1.1	8 - 27	25
tre1^{sctt} reconstruction (1)	1.0 ± 0.3	0 - 6	18.8 ± 1.3	6 - 32	19.8 ± 1.4	6 - 33	25
tre1^{sctt} reconstruction (2)	1.1 ± 0.3	0 - 5	14.0 ± 1.2	3 - 27	15.1 ± 1.2	3 - 27	25
AY ILI ACH (1)	0.7 ± 0.2	0 - 4	18.8 ± 1.2	10 - 35	19.5 ± 1.3	10 - 37	25
AY ILI ACH (2)	1.0 ± 0.3	0 - 5	17.5 ± 1.1	1 - 25	18.5 ± 1.1	3 - 28	25
RA ILI ACH (1)	12.5 ± 0.8	2 - 19	10.4 ± 1.1	0 - 25	22.9 ± 1.3	9 - 37	25
RA ILI ACH (2)	12.2 ± 1.3	0 - 23	11.8 ± 1.2	2 - 23	24.0 ± 1.3	15 - 38	25
RY AAA ACH	2.8 ± 0.4	0 - 4	15.1 ± 1.6	6 - 26	17.9 ± 1.5	10 - 27	15
	8.2 ± 1.5	5 - 21	14.2 ± 1.3	6 - 20	22.4 ± 1.3	15 - 27	10
RY AAA AAA (1)	2.0 ± 0.6	0 - 4	15.5 ± 1.6	12 - 21	17.5 ± 1.8	12 - 24	6
	7.1 ± 0.4	5 - 11	13.8 ± 1.1	5 - 24	21.0 ± 1.0	14 - 30	19
RY AAA AAA (2)	1.9 ± 0.3	0 - 4	12.4 ± 1.2	5 - 21	14.3 ± 1.3	7 - 23	15
	13.2 ± 2.0	5 - 21	10.1 ± 2.4	2 - 23	23.3 ± 1.3	17 - 28	10
RY AAA AAA recomb (1)	10.9 ± 0.7	5 - 16	15.0 ± 1.0	6 - 24	25.9 ± 1.4	16 - 40	25
RY AAA AAA recomb (2)	11.8 ± 0.8	4 - 18	14.1 ± 1.3	5 - 28	25.9 ± 1.4	13 - 37	25
RY AAA AAA recomb (3)	11.1 ± 0.9	1 - 19	15.6 ± 1.3	5 - 30	26.8 ± 1.4	16 - 42	25

Germ cell counts were performed on stages 15-16 embryos stained with a chicken anti-Vasa antibody to label the primordial germ cells. RYILIACH is the T^+G^+ vector described in Dahanukar et al., 2001.

ACKNOWLEDGEMENTS

I would like to thank a number of people who made this thesis possible. First, I would like to express gratitude to my advisor, Clark Coffman, for his guidance, advice and support throughout my graduate work. Through him I have grown as a scientist and as an independent researcher.

I wish to thank my husband, Spencer, for his never-ending encouragement, support and love. He has been my pillar for the past 8 years. I also want to thank my parents, Ken and Georgia Amdahl, for their love and support over the years and for raising me in a family that values education.

Lastly, I would like to thank past members of the Coffman Lab, especially Keri Davis and Angela Kamps, for their friendship and support through the challenging times of graduate school.